University of Alberta

An NMR investigation of Troponin I within the cardiac troponin complex and the effects on calcium induced cardiac muscle contraction: The 'I' in team

by

Darrin Anthony Lindhout ©

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

Edmonton, Alberta

Fall 2005

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

0-494-08681-5

Your file Votre référence
ISBN:
Our file Notre reference
ISBN:

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

"My religion consists of a humble admiration of the illimitable superior spirit who reveals himself in the slight details we are able to perceive with our frail and feeble minds."

Albert Einstein (1879 – 1955)

Dedicated to my loving grandparents:

Jim & Hilly Terpstra and John & Joanne Lindhout

ABSTRACT

Cardiac contraction is initiated via a calcium signaling pathway in which specific protein components within the sarcomere participate in a continual rhythmic cycle of contraction and relaxation. This thesis addresses the structural organization of the protein components of the thin filament with a focus on the protein-protein interactions made by and within the troponin complex during this signaling event. The troponin complex is composed of three proteins: C, I, and T. Troponin C is a calcium sensitive switch that monitors intracellular calcium concentrations, troponin I is responsible for the inhibition of contraction in the absence of calcium signaling, and troponin T is the protein component that anchors the troponin complex onto the thin filament. Multiple regions of troponin I have been shown to interact with troponin C, with several mechanisms proposed regarding the specifics of the altered protein-protein interactions occurring during a contraction cycle. Of these interactions none is more important than the repositioning of the inhibitory region of troponin I within the troponin complex.

The goal of this thesis is to expand the current knowledge of the interactions of the inhibitory region of troponin I within the troponin complex. Through the utilization of cloning, electrostatic and mutagenesis techniques, this thesis directly addresses physiological perspectives such as the role of phosphorylation as an important regulator of cardiac contraction, and the occurrence of diseased states such as familial hypertrophic cardiomyopathy due to mutational events. Structural, dynamic, and thermodynamic studies of troponin I have been pursued using modern high-resolution, multinuclear, and multidimensional nuclear magnetic resonance spectroscopy. The delicate electrostatic

balance occurring within this protein complex, deemed important to the functionality of the contraction cycle, is presented in the context of recently published biophysical and structural data.

ACKNOWLEDGEMENTS

First and foremost, I wish to express my most sincere gratitude to my supervisor and mentor; Professor Brian Douglas Sykes, for your constant encouragement, support, guidance, and infinite patience during both my undergraduate and graduate time in your laboratory. Your free thinking attitude, coupled with a sharp wit, keen sense of humour and an innate ability to foresee potential project problems has helped me to persevere through both hard and good times. In addition to being an extremely talented researcher, your ability to provide guidance as both a mentor and a friend has been invaluable, equipping me with all the necessary skills that I will need for all future endeavors. A huge thank you to your wife Nancy, who has been like a second mother to me: always offering advice on life and steering me in the right direction. My only regret that I have is that I cannot stay longer.

Many thanks to my supervisory committee which in addition to Brian D. Sykes, consisted of Drs. Mark Glover and Leo Spyracopoulos. Particular thanks to Dr. Spyracopoulos, who back in his post-doctoral days within the Sykes laboratory spent time with a naïve, scared undergraduate and taught me the ways of protein NMR with a smile and patience. Leo: many thanks for your computer expertise, hard work ethic, and overall passion for scientific research. I must also thank Dr. Ronald MacElhaney: student graduate coordinator, for his sometime bothersome but necessary reminders to regularly schedule committee meeting and meet the departmental requirements for graduation in a reasonable amount of time.

To Dr. Monica Li: I cannot underestimate the amount of time regarding both physiological knowledge and NMR expertise you have equated upon me over the years. I would never have succeeded with my project without your constant support, supervision and timely suggestions on where next to go. Thanks to Angela Thiessen, Linda Saltibus and Melissa Rakovszky; the technicians in my life who put up with me for over 5 years and did all my ordering, and assisted in cloning and protein purifications. Who could forget David Corson? The man who taught me everything I wanted to know about protein expression/purification but was afraid to ask. Your knowledge far outweighs your years of experience, don't ever change Dave! Thanks to Gerry McQuaid and Jeff DeVries for their tireless maintenance of our fleet of NMR spectrometers, and to Gerry's lovely wife Maureen who kept me well fed during the Christmas holiday seasons with shortbread baking.

My many thanks to my fellow graduate students and post doctoral fellows in the laboratory through the years. Your friendship was very much appreciated. Specific thanks goes out to Dr. Pascal Mercier and Xu Wang, both of whom are to be commended for their patience and teaching in computer programming and writing of scripts, never succumbing to my seeming endless ability to forget what you just told me. Both of you always made my problems a priority and never lost your cool when I asked how to do something for the 15th time (sometimes all in the same day). To 'team troponin': Xu Wang, Monica Li, Tharin Blumenschein, Pascal Mercier, Ryan Hoffman, Leo Spyracopoulos, Grant Gay, Ryan McKay and Stéphane Gagné, thank you for all your support, helpful discussions, trouble shooting and friendship. I would also like to thank the following laboratory personnel for their friendship throughout the year: Steffen Graether, Jan Rainey, J.Y. Suh, David Kaiser, Vincent Raussens, Leigh Willard, and Dean Schieve.

In particular I would like to thank both Erik Saude and Robert Boyko for their friendship throughout the years. Erik: you helped me to look to the lighter side of science and realize that it wasn't always as bad as I might have thought it could be. Your constant dedication to pursue only the best of both science and physique has taught me a lot. You're the man!!.....no matter what everyone says.......*it's all just a metaphor anyways.* As for Robert Boyko, I am not too sure whether to thank you or curse you for getting me into long distance running, which has now laid claim to most of my free time and energy. You're an amazing person Robert with computer skills that are second to none. I will always have fond memories of our weekly dart games at Avenue Pizza, when you as usual would lay claim to whatever pride I might have possessed during that week.

I must also thank the friendly and knowledgeable staff at the National High Field Nuclear Magnetic Resonance Centre (NANUC) for setup and expertise during ultra-high field acquisitions of my protein samples, as well for letting me borrow a desk for a few months to write up my thesis. My thanks goes out to Stephane Gagne, Ryan McKay, Bruce Lix, Deryck Webb and Dipika Parmar for all your collective efforts. I must also thank the departmental support staff present for help with networking issues and secretarial matters. Thanks to: Ioannis Bastiolas, Dean Schieve, Roger Bradley, Perry D'Obrenan, Susan Smith, Susan Henry, Marion Benedict, and all the secretaries within the Department of Biochemistry. Special thanks also goes out to the boys at the Institute for Biomolecular Design: Paul Semchuck and Lorne Burke for their HPLC and MALDI-TOF mass spectrometry expertise.

To my friends that I have made during the long trip from start to finish, thank you for all the support and friendships. Vic Drover, Daisy Sahoo, Jodi Carter, Robert Ryan, Luc Berthiaume, Audric Moses, Russ Watts, Erik Labonte, Trent Letcher, Les Grad, Sam Szeto, Jonathan Lacasse, Kristen Conn, Tyler Muri, Colleen Schur and countless others are to be congratulated for putting up me and my antics.

How can I possibly put into words the absolute awe that I possess for my girlfriend Leanne? You're support is unquestionable and your patience has no equal. Your smile always brightens my day and makes all my troubles go away. A very special thanks to my family: Tony & Sharon Lindhout, Roland & Gerda Goulet and Cristine & Mark Gibson, for their constant support: both financially and mentally. Thanks to everyone!

Special monetary thanks goes to the Alberta Heritage Foundation for Medical Research for supplying myself with a generous studentship that has alleviated much of the financial concerns that presented during my graduate career. I also thank the Canadian Institution for Health Research and the Canadian Heart and Stroke Foundation for a high level of laboratory funding.

TABLE OF CONTENTS

Page

CHAPTER I – Cardiac muscle and the troponin complex

CHAPTER OVERVIEW 1

STRUCTURAL ORGANIZATION OF CARDIAC TISSUE 3

STRUCTURAL DIFFERENCES BETWEEN MUSCLE TISSUES 10

THE MECHANISM OF MUSCLE CONTRACTION 11

TROPONIN C – CALCIUM CONTROL OF MUSCULAR CONTRACTION 16

TROPONIN T – TROPOMYOSIN BINDING AND CALCIUM SENSITIVITY 19

TROPONIN I – INHIBITION OF CARDIAC CONTRACTION 20

THE INHIBITORY REGION OF TROPONIN I 23

TROPONIN I – MUTATION AND CARDIOMYOPATHY 24

TROPONIN I – POST-TRANSLATIONAL MODIFICATION 25

TROPONIN I – STRUCTURAL INSIGHTS FROM A CRYSTAL STRUCTURE 27

THESIS GOALS AND PROJECT RATIONALE 28

- *Specific Aim #1* – Production of labeled inhibitory region for NMR structural studies 29
- *Specific Aim #2* – Solve the NMR solution structure of the bound inhibitory region 30
- *Specific Aim #3* – Electrostatic and kinetic consequences on complex formation 30
- *Specific Aim #4* – The effects of cRP40 and the inhibitory region on binding to the C-lobe 30
- *Specific Aim #5* – Interactions of cTnI with the N-lobe of cTnC 31
- *Specific Aim #6* – NMR studies of the full length troponin complex 31
- *Specific Aim #7* – Correlating structural studies with functional studies 32

REFERENCES 33

CHAPTER II – Design of a peptide expression system for NMR studies of the troponin system

OVERVIEW 37
INTRODUCTION 38
EXPERIMENTAL PROCEDURES
II-A. GEV-1 vector and primer design 40
II-B. Sequencing of GEV1 fusion plasmids 43
II-C. Test Expression of GB1 fusion proteins 43
II-D. Labeling of GB1 fusion proteins cIp-RR-20 and TRTK-12 44
II-E. Purification of GB1 fusion proteins 44
II-F. CNBr cleavage of GB1 fusion proteins and purification of peptides 45
II-G. NMR spectroscopy 45
RESULTS 45
DISCUSSION 52
REFERENCES 55

CHAPTER III – The inhibitory region of cardiac troponin I: structure and function

OVERVIEW 57
INTRODUCTION 58
EXPERIMENTAL PROCEDURES
III-A. Preparation of cCTnC protein 63
III-B. Preparation of cIp protein 64
III-C. Isotope labeling of cCTnC and cIp proteins 64
III-D. ^{1}H-NMR structural studies on unbound cIp protein 65
III-E. Titration of synthetic cIp peptide with ^{15}N/^{13}C-cCTnC•2Ca^{2+} 65
III-F. Titration of ^{15}N-cCTnC•2Ca^{2+} with synthetic cIp peptide 65
III-G. Titration of ^{15}N-cIp with cCTnC•2Ca^{2+} 66
III-H. Structural studies on cCTnC•2Ca^{2+}•cIp complex 66
III-I. NMR Spectroscopy 67
III-J. Data processing and peak calibration 67
III-K. Structural calculations for binary complex cCTnC•2Ca^{2+}•cIp 67

III-L. Backbone amide ^{15}N relaxation measurements of cCTnC•2Ca^{2+}•cIp complex 68
III-M. Coordinates 69

RESULTS

III-A. Structural studies of unbound synthetic ^{1}H-cIp 69
III-B. Induced chemical shift changes to ^{1}H-cIp upon binding to the C-lobe 71
III-C. Titration of ^{15}N-cIp with cCTnC•2Ca^{2+} 72
III-D. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp 74
III-E. Overall Structure of the cCTnC•2Ca^{2+}•cIp complex 79
III-F. Structure of cCTnC in the complex 83
III-G. Structure of cIp in the complex 85
III-H. Backbone amide ^{15}N-NMR relaxation studies of cCTnC 89
III-I. Backbone amide ^{15}N-NMR relaxation studies of ^{15}N-cIp 91
DISCUSSION 92
REFERENCES 104

CHAPTER IV – The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C

OVERVIEW 108
INTRODUCTION 109

EXPERIMENTAL PROCEDURES

IV-A. Preparation of cCTnC protein and synthetic peptides 112
IV-B. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-K139G 112
VI-C. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-R140G 113
VI-D. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-R144G 114
VI-E. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-R145G 114
VI-F. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-T142Phos 114
VI-G. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-R144G-T142Phos 115
VI-H. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-G4 (K139G, R140G, R144G, R145G) 115
VI-I. NMR Spectroscopy 116

RESULTS

IV-A. Titration of cIp-G4 peptide with the C-lobe of cTnC 117
IV-B. The consequences of cIp mutation on cCTnC binding 117
IV-C. Distance effects on electrostatic attractive forces 122
DISCUSSION 127
REFERENCES 132

CHAPTER V – Interactions of cRP40 and the inhibitory region of troponin I with troponin C

OVERVIEW 134
INTRODUCTION 135

EXPERIMENTAL PROCEDURES

V-A. Preparation of cCTnC protein and synthetic peptides 137
V-B. Competition binding studies of cIp and cRP40 peptides for ^{15}N-cCTnC•2Ca^{2+} 138
V-C. Titration of ^{15}N-cTnC•3Ca^{2+} with cRP40 and cRP40-S41Phos/S43Phos 139
V-D. NMR Spectroscopy 139
V-E. ^{15}N-NMR T1 and T2 relaxation parameters 140

RESULTS

V-A Binding studies of cIp and cRP40 on C-lobe of cTnC 140
V-B Effect of Ser41/Ser43 phosphorylation on the interaction of cRP40 with cTnC•3Ca^{2+} 145
DISCUSSION 148
REFERENCES 154

CHAPTER VI – Troponin interactions with the N-lobe of cardiac troponin C

OVERVIEW 156
INTRODUCTION 157

EXPERIMENTAL PROCEDURES

VI-A. Preparation of cNTnC protein and synthetic peptides 160

VI-B. Titration of cIp with the N-lobe of cTnC 161
VI-C. Titration of cSp with cNTnC•cIp complex 161
VI-D. Titration of cSp with the N-lobe of cTnC 162
VI-E. Titration of cIp with cNTnC•cSp complex 162
VI-F. Titration of ^{15}N-cNTnC with cSp, cSp-S149Phos, cSp-R161W, and cSp-R161W-S149Phos 162
VI-G. NMR spectroscopy 163

RESULTS

VI-A. Effects of cIp and cSp binding on the N-lobe of cTnC 163
VI-B. Effect of Ser149 phosphorylation and R161W mutation on the interaction of cSp with cNTnC•Ca^{2+} 168
DISCUSSION 173
REFERENCES 178

CHAPTER VII – Troponin I interactions within a 38kDa Tn complex

OVERVIEW 180
INTRODUCTION 181

EXPERIMENTAL PROCEDURES

VII-A. Cloning, expression and purification of cTnT$_{197-298}$ 183
VII-B. Cloning, expression and purification of cTnI$_{34-163}$ 184
VII-C. Sequencing cTnT$_{197-298}$ and cTnI$_{34-163}$ pET3a vectors 185
VII-D. Synthesis of cTnT$_{236-271}$ peptide 185
VII-E. Production of a high molecular weight 45.4kDa troponin complex 186
VII-F. Verification of a 45.4kDa troponin complex by electrophoresis 186
VII-G. Production of a 38.8kDa troponin complex 187
VII-H. NMR spectroscopy 187
VII-I. ^{15}N-NMR T1 and T2 relaxation parameters 188

RESULTS

VII-A. Choice of peptide cTnT$_{226-271}$ for use in the troponin complex 190
VII-B. NMR studies of cTnC within a high molecular weight troponin complex 192

VII-C. NMR studies of $cTnI_{34-163}$ within a high molecular weight troponin complex 193
VII-D. Comparison of skeletal and cardiac TnI within the troponin complex 197
DISCUSSION 200
REFERENCES 205

CHAPTER VIII – Conclusion

OVERVIEW 207
SUMMARY 207
DISCUSSION 209
REFERENCES 217

LIST OF TABLES

Page

CHAPTER I - Cardiac muscle and the troponin complex

Table I-I. Interhelical angles of N-lobe. 4

CHAPTER III – The inhibitory region of cardiac troponin I: structure and function

Table III-I. NMR spectra acquired and experimental conditions used for cIp peptide 68
Table III-II. NMR spectra acquired and experimental conditions used to obtain assignments and NOE restraints 80
Table III-III. Structural statistics of the family of the 30 structures calculated 83
Table III-IV. Interhelical angles of various EF hands 85

CHAPTER IV – The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C

Table IV-I. Dissociation constants of various cIp peptides binding to C-lobe of cTnC 119
Table IV-II. Thermodynamic data of cIp peptides binding to C-lobe of cTnC 125

CHAPTER VII – Troponin I interactions within a 38kDa Tn complex

Table VII-I. NMR spectra acquired and experimental conditions 189

APPENDIX B – Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na$^+$/H$^+$ exchanger

Table B-1. Summary of subcellular localization in AP-1 cells of wild type (wNHE1), cysteineless NHE1 (cNHE1), and of TM IV Na$^+$/H$^+$ exchanger mutants. 243
Table B-2. Structural statistics for the final ensemble of 600 structures out of the 1000 calculated for $CD_3OH:CDCl_3:H_2O$. 249

Table B-3. RMSD for C^{α} and for all heavy atoms of the 600 lowest energy structures of the 1000 member ensemble of the TM IV peptide superposed over each convergent stretch and given relative to the average structure for that superposition. 253

APPENDIX E – Using lanthanide ions to align troponin complexes in solution: Order of lanthanide occupancy in cardiac troponin C

Table E-1. The active binding loops of (A) chicken cardiac and (B) rabbit skeletal troponin C 297

APPENDIX F – NMR solution structure of a highly stable de novo heterodimeric coiled-coil

Table F-1. Structural Statistics of the family of the 20 structures calculated 309
Table F-2. Distances of Salt-Bridging pairs within IAAL-E3/K3 coiled-coil 318

LIST OF FIGURES

Page

CHAPTER I - Cardiac muscle and the troponin complex

Figure I-1. Architecture of muscle tissue. 4
Figure I-2. Anatomy of the myofibril. 5
Figure I-3. Maximal tension and the sliding filament hypothesis. 6
Figure I-4. Structure of the thick filament. 7
Figure I-5. Structure of the thin filament. 8
Figure I-6. Proteins of the sarcomere. 9
Figure I-7. Intercalating discs. 11
Figure I-8. Calcium release from the sarcoplasmic reticulum. 12
Figure I-9. Ca^{2+}-induced thin filament repositioning of tropomyosin. 14
Figure I-10. Mechanism of force generation in muscle. 15
Figure I-11. Structure of cardiac troponin C. 17
Figure I-12. Calcium binding to N-lobe. 19
Figure I-13. Cardiac troponin T variability. 20
Figure I-14. Calcium induced structural changes of the troponin complex. 21
Figure I-15. Structure of the 'core' troponin complex. 28

CHAPTER II - Design of a peptide expression system for NMR studies of the troponin system

Figure II-1. GEV1 vector map. 41
Figure II-2. Design of peptide DNA oligonucleotides for insertion into the GEV1 vector 42
Figure II-3. *Escherichia coli* K12 codon bias. 43
Figure II-4. GB1 fusion protein constructs. 46
Figure II-5. 8-22% gradient SDS-PAGE gel of expressed fusion proteins. 47
Figure II-6. Cyanogen Bromide cleavage of amide bond. 48
Figure II-7. Cyanogen Bromide cleavage of GB1 fusion protein. 49
Figure II-8. HPLC chromatogram of CNBr cleaved fusion protein. 50

Figure II-9. HSQC NMR spectra of cIp-RR-20 peptide. 52

CHAPTER III – The inhibitory region of cardiac troponin I: structure and function

Figure III-1. Unbound cIp in solution. 70

Figure III-2. Titration of ^{1}H-cIp with C-lobe. 71

Figure III-3. Titration of ^{15}N- cIp with ^{1}H-cCTnC. 73

Figure III-4. Titration of ^{15}N-cCTnC with ^{1}H-cIp. 74

Figure III-5. Comparison of induced chemical shift changes of cCTnC•Ca^{2+} upon the addition of cIp peptide. 76

Figure III-6. Chemical shift map of the C-lobe during cIp binding. 77

Figure III-7. Dissociation curve of cIp binding to the C-lobe. 78

Figure III-8. The structure of the bound inhibitory region with the C-lobe. 82

Figure III-9. ^{13}C-NOESY strip-plot of Val160 connectivities. 84

Figure III-10. Electrostatic surface map of binary complex cCTnC•2Ca^{2+}•cIp. 88

Figure III-11. C-lobe ^{15}N-amide relaxation measurements. 89

Figure III-12. C-lobe ^{15}N-amide relaxation measurements. 91

Figure III-13. Model of skeletal troponin I in complex with troponin C. 94

Figure III-14. Helical models of the inhibitory region of skeletal troponin I. 95

Figure III-15. Methionine solvent protection of C-lobe upon cIp binding. 96

Figure III-16. Lindhout model of inhibitory region bound to cardiac troponin. 98

Figure III-17. Crystal structure of the 'core' cardiac troponin 52kDa complex. 100

Figure III-18. Model of the 'core' cardiac troponin complex of the N-lobe. 101

CHAPTER IV – The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C

Figure IV-1. Mutagenesis of cIp peptides used to monitor the electrostatic binding effects of cIp to the C-lobe of cTnC. 117

Figure IV-2. Dissociation constants (K_D) for various cIp mutant peptides with the C-lobe of cTnC. 118

Figure IV-3. Induced C-lobe chemical shift changes by cIp peptides. 120

Figure IV-4. Mutation and phosphorylation of the inhibitory region. 121

Figure IV-5. Mutational analysis of the inhibitory peptide. 124

Figure IV-6. Linear relationship of free energy of binding (ΔG) versus potential energy (ΔE). 126

CHAPTER V – Interactions of RP40 and the inhibitory region of troponin I with the C-lobe of cTnC

Figure V-1. Steric clash of the inhibitory and cRP40 regions of cTnI. 140

Figure V-2. ^{15}N-HSQC plots of ^{15}N-cCTnC•Ca^{2+} titration with cIp and cRP40. 141

Figure V-3. ^{15}N-HSQC plots of ^{15}N-cCTnC•Ca^{2+} titration with cRP40 and cIp. 142

Figure V-4. ^{15}N-NMR backbone relaxation of the C-lobe of cTnC in complex with cTnI peptides. 143

Figure V-5. Linear relationship of ^{15}N-NMR R2 relaxation values versus molecular weight for various troponin protein complexes. 144

Figure V-6. Titration of ^{15}N-cTnC•3Ca^{2+} with cRP40 and cRP40-Ser41Phos/Ser43Phos. 145

Figure V-7. Ribbon and surface representations of the structure of cCTnC•2Ca^{2+}•cRP40. 147

Figure V-8. 'Lindhout' model of the 'core' cardiac troponin complex. 150

Figure V-9. Crystal structure of the skeletal troponin complex. 151

CHAPTER VI – Troponin interactions with the N-lobe of cardiac troponin C

Figure VI-1. Ribbon representation of cNTnC in (A) the calcium bound state and (B) the cSp bound state. 157

Figure VI-2. ^{15}N-HSQC spectra of cNTnC-A$_{cys}$. 163

Figure VI-3. cIp-cSp titration of ^{15}N-cNTnC-A$_{cys}$. 164

Figure VI-4. cSp-cIp titration of ^{15}N-cNTnC-A$_{cys}$. 165

Figure VI-5. Ribbon diagram of calcium saturated cNTnC bound to cSp. 166

Figure VI-6. N-lobe induced chemical shifts upon cIp binding. 167

Figure VI-7. Chemical shift map of the effects of cIp binding on the surface of calcium saturated cNTnC. 168

Figure VI-8. Titration of ^{15}N-cNTnC•Ca^{2+} with cSp-S149Phos, cSp-R161W, and cSp-R161W-S149Phos. 169
Figure VI-9. Chemical shift changes of N-lobe upon cSp binding. 171
Figure VI-10. Positioning of Arg147 in the troponin complex. 174

CHAPTER VII – Troponin I interactions within a 38kDa Tn complex

Figure VII-1. Coiled-coil interaction of cTnT and cTnI subunits. 190
Figure VII-2. Native PAGE gel strip plots of a 38.8kDa ternary troponin complex. 191
Figure VII-3. {^{1}H, ^{15}N}-HSQC of ^{15}N-cTnC binding to troponin. 192
Figure VII-4. {^{1}H, ^{15}N}-HSQC of ^{15}N-cTnC binding to troponin subunits – Gly110 and Gly146. 193
Figure VII-5. Spectral dispersion of the cTnI$_{34-163}$ at 800MHz. 194
Figure VII-6. Amide dispersion comparison of the inhibitory region and cTnI$_{34-163}$. 195
Figure VII-7. Example of cTnI$_{34-163}$ amide connectivities. 196
Figure VII-8. Comparison of skeletal and cardiac TnI in the troponin complex. 197
Figure VII-9. ^{15}N-R2 versus molecular weight for cTnI$_{34-163}$. 198
Figure VII-10. ^{15}N-T2 NMR relaxation values for cTnI$_{34-163}$ in the troponin complex. 199
Figure VII-11. Comparison of cardiac and skeletal crystal structures. 202
Figure VII-12. Domain orientations of the troponin complex. 203

APPENDIX A – Establishment of a high yield recombinant peptide expression System

Figure A-1. His-tag fragment. 225
Figure A-2. cIp-RR-20 peptide fragment. 226
Figure A-3. cIp-RR-20 peptide fragment. 227
Figure A-4. GB1 peptide fragment. 228
Figure A-5. GB1 peptide fragment. 229

APPENDIX B – Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger

Figure B-1. Topological model of the NHE1 isoform of the Na^+/H^+ exchanger. 239
Figure B-2. Rate of recovery. 240
Figure B-3. Western blot analysis of cell extracts from control and stably transfected AP-1 cells. 241
Figure B-4. Effects of MTSET and MTSES on activity of cNHE1 and single cysteine mutant NHE1 containing cell lines. 242
Figure B-5. Example of a Western blot for determination of localization of the Na^+/H^+ exchanger. 244
Figure B-6. Summary of observed NOE's. 246
Figure B-7. The NMR Structure of TM IV. 254

APPENDIX E – Using lanthanide ions to align troponin complexes in solution: Order of lanthanide occupancy in cardiac troponin C

Figure E-1. HSQC lanthanide titration of cTnC. 288
Figure E-2. Ce^{3+} titration of cTnC. 289
Figure E-3. cTnC HSQC. 290
Figure E-4. Ce^{3+} titration of sTnC. 291
Figure E-5. RDC measurements of bound inhibitory region. 294
Figure E-6. RDC plot of bound inhibitory region. 295

APPENDIX F – NMR solution structure of a highly stable de novo heterodimeric coiled-coil

Figure F-1. Sequence and diagrammatic representation of the IAAL-E3/K3 protein complex. 310
Figure F-2. Amide proton resonance region of the ^{1}H-NMR spectrum. 311
Figure F-3. Amide proton resonance region of the ^{1}H-NMR spectrum. 313
Figure F-4. Structure of the heterodimeric IAAL-E3/K3 coiled-coil. 314
Figure F-5. The structure of the heterodimeric IAAL-E3/K3 domain. 316

LIST OF ABBREVIATIONS

TnC – troponin C

TnI – troponin I

TnT – troponin T

cTnC – cardiac troponin C

cTnI – cardiac troponin I

cTnT – cardiac troponin T

NMR – Nuclear Magnetic Resonance

SR – Sarcoplasmic Reticulum

CaRC – Calcium Release Channel

P_i – inorganic phosphate

cRP40 – N-terminal 47-residue region cTnI$_{33\text{-}80}$

cIp – inhibitory region cTnI$_{128\text{-}147}$

cSp – switch region cTnI$_{147\text{-}163}$

cAp – actin-binding region of cTnI$_{164\text{-}186}$

HC – Hypertrophic Cardiomyopathy

FHC – Familial Hypertrophic Cardiomyopathy

PKA – Protein Kinase A

PKC – Protein Kinase C

PAK – p21-activated kinase

bpm – beats per minute

RH – Regulatory Head

HSQC – Heteronuclear Single Quantum Coherence

NANUC – National High Field Nuclear Magnetic Resonance Center

TROSY – Transverse Relaxation Optimized Spectroscopy

hIAPP – islet amyloid polypeptide

GB1 – protein G from the *Streptococcal* B1 immunoglobulin-binding domain

capZα1 – actin capping protein

CNBr – Cyanogen Bromide

HPLC – High Pressure Liquid Chromatography

IPTG – Isopropyl β-D-thiogalactoside
SDS-PAGE – Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
2xYT – Yeast/Tryptone media (2X concentrate)
M9 – Minimal Media containing phosphate buffer
ddH_2O – distilled, deionized water
MALDI-TOF - Matrix Assisted Laser Desorption Ionization-Time of Flight
MW – Molecular Weight
MS – Mass Spectrometry
hNHE1 - human sodium-proton exchanger isoform 1
ppm – parts per million
hSer - homoserine
cCTnC – cardiac TnC (C-lobe)
sTnI – skeletal Troponin I
sTnC – skeletal Troponin C
sTnT – skeletal Troponin T
sRP40 - N-terminal 47-residue region $sTnI_{1-47}$
RDC – Residual Dipolar Coupling
nOe – nuclear Overhauser effect
K_D – Dissociation Constant
r.m.s.d. – Root Mean Squared Deviation
sCTnC – skeletal TnC (C-lobe)
EPR – Electron Paramagnetic Resonance
IUPAC – International Union of Pure and Applied Chemicals
sNTnC – skeletal TnC (N-lobe)
cNTnC – cardiac TnC (N-lobe)
SANS – Small Angle Neutron Scattering
TFE – Trifluoro ethanol
LTnH – *Lethocerus* troponin H

LIST OF APPENDICES

Page

APPENDIX A – Establishment of a high yield recombinant peptide expression system

Appendix A-1. Codon usage in Escherichia coli K12 219

Appendix A-2. Sequencing results of GEV1 peptide expression vectors 221

Appendix A-3. Sequencing results of pET3a cTnI and cTnT expression vectors 223

Appendix A-4. MALDI-TOF mass spectrometry results of HPLC purified peptide 225

APPENDIX B – Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger

OVERVIEW 230

ABBREVIATIONS 231

INTRODUCTION 231

EXPERIMENTAL PROCEDURES

B.1 Materials 233

B.2 Site-directed mutagenesis 233

B.3 Cell culture and stable transfection 233

B.4 SDS-PAGE and immunoblotting 234

B.5 Cell surface expression 234

B.6 Na^+/H^+ exchange activity and sulfhydryl reactive reagent inhibition 234

B.7 Production and purification of TM IV 235

B.8 NMR spectroscopy and structure calculation 237

B.9 Coordinates 238

RESULTS

B.1 Activity and expression of NHE1 mutants 239

B.2 Subcellular localization of the mutant and wild-type Na^+/H^+ exchangers 244

B.3 Production and characterization of TM IV 245
B.4 Choice of NMR conditions 246
B.5 Resonance assignment 247
B.6 Structure calculation 248

DISCUSSION

B.1 Functional analysis of TM IV 249
B.2 Structural analysis of TM IV 251
B.2 Correlating structural and functional data 256
CONCLUSIONS 257
REFERENCES 258

APPENDIX C – Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na$^+$/H$^+$ exchanger - chemical shift and procheck structural analysis

Appendix C-1. Assigned ^{1}H and ^{15}N chemical shifts of TMIV of hNHE1 260
Appendix C-2. Procheck analysis of TMIV of hNHE1 262

APPENDIX D – The inhibitory region of cardiac troponin I: structure and function - chemical shift and procheck structural analysis

Appendix D-1. Assigned ^{1}H, ^{13}C and ^{15}N chemical shifts of unbound cIp in solution 266
Appendix D-2. ^{1}H, ^{13}C and ^{15}N chemical shifts of cCTnC•2Ca2$^+$•cIp binary complex 267
Appendix D-3. Procheck analysis of cCTnC•2Ca2$^+$•cIp binary complex 273

APPENDIX E – Using lanthanide ions to align troponin complexes in solution: Order of lanthanide occupancy in cardiac troponin C

OVERVIEW 280
ABBREVIATIONS 281
INTRODUCTION 281

EXPERIMENTAL PROCEDURES

E.1 Sample Preparation 284

E.2 Lanthanide titrations of cTnC 285

E.3 Lanthanide titration of sTnC 285

E.4 Yb^{3+} Titration of $cTnC\bullet3Ca^{2+}\bullet{}^{15}N\text{-}cTnI_{128-147}$ 286

E.5 NMR spectroscopy 286

RESULTS

E.1 Lanthanide titration of Ca^{2+}-saturated cTnC 287

E.2 Lanthanide titration of Apo-cTnC 290

E.3 Lanthanide titration of Ca2+-saturated sTnC 291

E.4 Orientation of $cTnC\bullet Yb^{3+}\bullet2Ca^{2+}$ and RDC measurements of bound $^{15}N\text{-}cTnI_{128-147}$ 292

DISCUSSION 295

CONCLUSIONS 299

REFERENCES 300

APPENDIX F – NMR solution structure of a highly stable de novo heterodimeric coiled-coil

OVERVIEW 303

ABBREVIATIONS 304

INTRODUCTION 304

EXPERIMENTAL PROCEDURES

F.1 Synthesis 307

F.2 NMR Spectroscopy 307

F.3 Structure Determination 308

F.4 Coordinates 308

RESULTS AND DISCUSSION 309

REFERENCES 320

APPENDIX G – NMR solution structure of a highly stable de novo heterodimeric coiled-coil – chemical shift and procheck structural analysis

Appendix G-1. Assigned ^{1}H and ^{15}N chemical shifts of EK coiled-coil 322
Appendix G-2. Procheck analysis of EK coiled-coil 324

CHAPTER I

Cardiac muscle and the troponin complex

CHAPTER OVERVIEW

This thesis addresses structural, thermodynamic and dynamic properties of the troponin complex found within cardiac tissue. The troponin complex is the protein component directly responsible for initiation of cardiac contraction. The troponin complex is composed of 3 distinct subunits referred to as troponin C (cTnC), troponin I (cTnI), and troponin T (cTnT). All three subunits play critical roles in the calcium initiation signal to begin cardiac contraction. *This thesis will focus on the cTnI subunit*: a protein that is responsible for inhibiting contraction during cardiac relaxation. Particular emphasis is placed on a small sub-domain of cTnI referred to as the inhibitory region, a highly basic 20 amino acid domain that is responsible for interactions with actin, troponin cTnT and cTnC. Mutation events within cTnI have been linked to diseased states and impaired cardiac function. Cellular regulation of the contraction/relaxation process can be altered in response to stress events via phosphorylation at specific sites throughout cTnI. High resolution, multinuclear and multidimensional nuclear magnetic resonance (NMR) spectroscopy was employed to obtain *in vitro* data regarding various subunits of cTnI and their respective binding partners: cTnC and cTnT. Mutational events and post-translational modification by phosphorylation within cTnI were investigated to ascertain any structural consequences, and will be discussed in relation to the *in vivo* physiologic responses attributed with these states. The NMR data elucidated from this body of work is central to the understanding of the molecular architecture responsible for muscle contraction.

This chapter concentrates on the general physiology and structural aspects of the cardiac muscle machinery, building from the macro-molecular environment of cardiac tissue to the micro-molecular environment of protein-protein interactions within the thick and thin filaments. Specific emphasis is placed on the importance of the interactions between the cTnC, cTnI and cTnT subunits, and on the rationale for which this research was motivated. This chapter is not intended to be an exhaustive reiteration of the large amount of information on muscle contraction that is currently available from detailed reviews in the literature [(*1-13*) and references within]. All data presented in subsequent chapters deals exclusively with the cardiac isoform(s) of the troponin complex. Inferences to published data regarding isoform differences in relation to the skeletal system will be examined; however, they will not be discussed in a high level of detail. This chapter will not review any of the NMR aspects associated with sample preparation, data collection, pulse sequences, data processing, peak and chemical shift assignments and/or structure generation. Likewise, no review will be offered regarding specific cloning, expression, and/or purification strategies of the proteins and complexes studied, these will be presented in subsequent chapters.

STRUCTURAL ORGANIZATION OF CARDIAC TISSUE

Cardiac muscle contraction and relaxation are biological processes carried out by an efficient and finely tuned molecular machine of highly synchronized protein-protein interactions with integrated activity. An important challenge in the study of cardiac muscle physiology is to gain understandings regarding alterations at the molecular level to the complexity of biological contraction and relaxation. Through the detailed atomic knowledge of all protein-protein interactions occurring during muscular activity, therapies and rational drug design can be developed to address injured tissues with altered activities, arising from genetic mutations and/or diseased states. This section presents an overall picture of the architecture within cardiac muscle fibers, and introduces the nature and function of the main protein components within cardiac myofibrils.

The heart is a muscular organ of the circulatory system that constantly pumps blood throughout the body. A healthy human heart is approximately the size of a clenched fist, composed of muscle tissue that is very strong, and is able to contract and relax rhythmically throughout a person's lifetime. The heart has four separate compartments or chambers. The upper chamber on each side of the heart is called an atrium, which receives and collects the blood coming into the heart. The atrium then delivers blood to a powerful lower chamber called a ventricle, which pumps blood away from the heart through powerful, rhythmic contractions. Cardiac muscle cells are generally referred to as *cardiocytes* and/or *cardiac myocytes*, and are relatively small; averaging 10–20*u*m in diameter and 50–100*u*m in length (see Figure I-1). Within cardiocytes, contractive forces are generated through an abundance of organized networks of filamentous fibers, referred to as *myofibrils*, which are arrayed in a parallel fashion. As viewed by electron microscopy (see Figure I-2), myofibrils possess highly repeating, ordered arrays of light (lesser electron density) and dark bands (greater electron density). The alternating isotropic light bands (I-band) and anisotropic dark bands (A-band) display darker regions at their centers due to the presence of both Z disks (I-band) and M disks (A-band), which primarily serve as anchoring units for the two main types of muscular filaments: the *thin* and *thick filaments*. The A-band consists of

150Å diameter thick filaments and the I-band consists of 70Å diameter thin filaments, which are linked together by cross-bridges where they overlap. The region that separates two adjacent Z-disks is referred to as the *sarcomere* and constitutes the basic contractile unit of the myofibril.

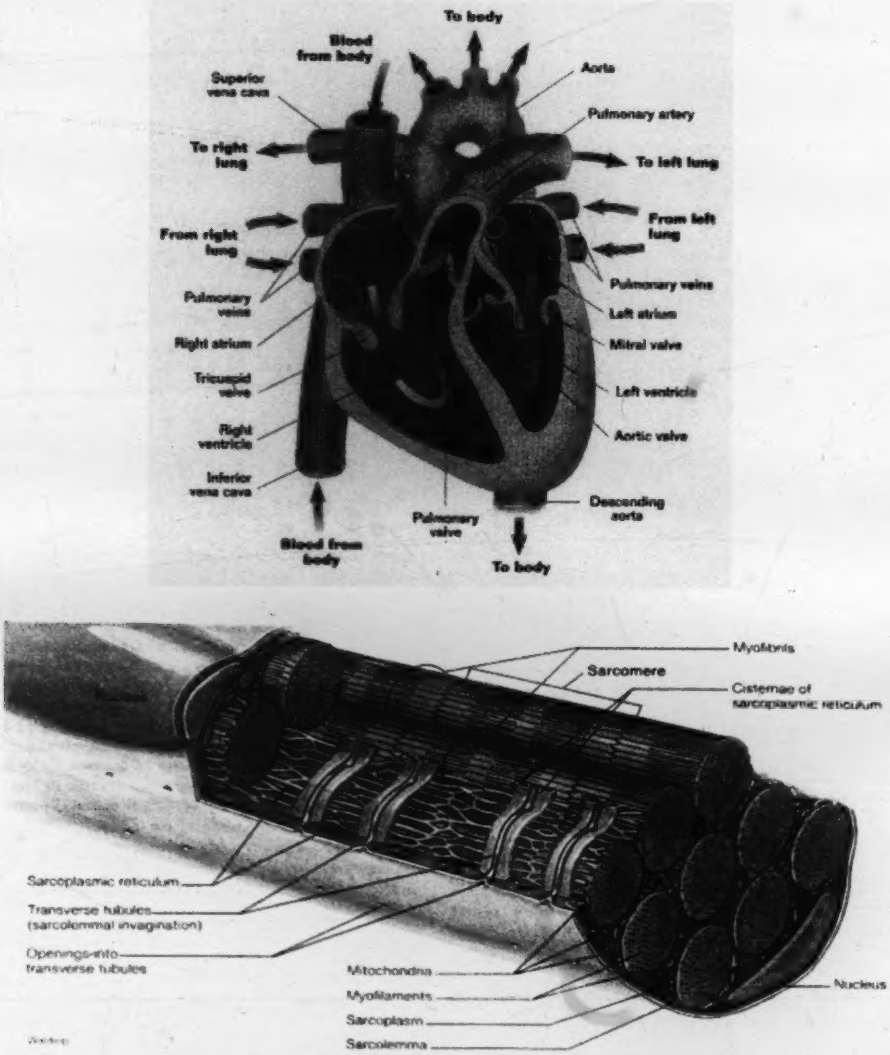

To body
Blood from body
Superior vena cava
Aorta
Pulmonary artery
To right lung
To left lung
From right lung
From left lung
Pulmonary veins
Pulmonary veins
Left atrium
Right atrium
Mitral valve
Tricuspid valve
Left ventricle
Right ventricle
Aortic valve
Inferior vena cava
Descending aorta
Pulmonary valve
Blood from body
To body
Myofibrils
Sarcomere
Cisternae of sarcoplasmic reticulum
Sarcoplasmic reticulum
Transverse tubules (sarcolemmal invagination)
Openings into transverse tubules
Mitochondria
Myofilaments
Sarcoplasm
Sarcolemma
Nucleus

Figure 1-1. **Architecture of muscle tissue**. Diagram of the front side of the human heart with specified arteries and veins [top]. Specific cellular components and organelles within cardiac tissues are as labeled, indicating the highly organized and intertwined cellular frame required for muscle contraction [bottom]. Figures adapted from (*14-16*).

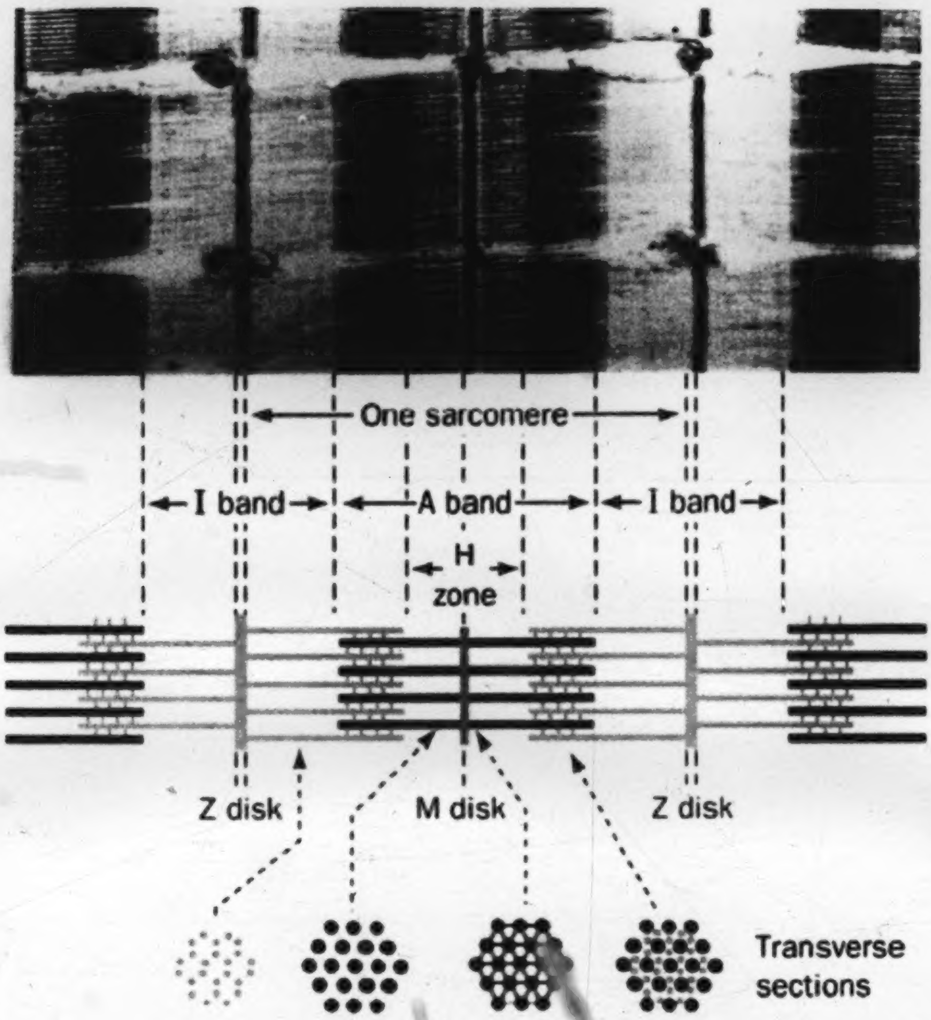


Figure I-2. **Anatomy of the myofibril**. An electron micrograph depicting a cross-sectional array of three continuous myofibrils, which are separated by horizontal white gaps [top]. The major features of the myofibril are as indicated as a cartoon [middle], where I-bands are areas of lesser electron density, A-bands are areas of greater electron density, with the Z and M disks anchoring the respective thin (pink) and thick (purple) filaments. The H-zone contains only M-disk and thick filament. The myofibril's functional unit of one sarcomeric length is indicated as the distance between neighboring Z-disks. Transverse sections displaying a cross-sectional cartoon representation of protein interactions, corresponding to electron micrograph densities are as indicated [bottom]. This figure and legend were both adapted from (*17*).

In cardiac muscle, maximal force generation is obtained at a sarcomere length of 2.2*u*m. The heart normally operates at sarcomere lengths less than this optimal length (usually 1.8*u*m – 2.0*u*m), therefore increases in sarcomere length result in an increase in tension development by the sarcomere. For various structural and mechanical reasons, the sarcomere length in the cardiac myocytes does not normally exceed 2.2*u*m (*18*). The length-tension diagram shows that as preload increases (Figure I-3A), there is an increase in active tension up to a maximal limit. The maximal active tension corresponds in cardiac muscle to a sarcomere length of 2.2*u*m. Cardiac muscle (unlike skeletal muscle) does not display a descending limb on the active tension curved because the greater stiffness of cardiac muscle normally prevents its sarcomeres from being stretched beyond 2.2*u*m. The contraction within the sarcomere results from reductions in the lengths of both the I-band and the H-zone (refer to Figure I-2). These observations yielded the hypothesis of a 'sliding filament model' over 50 years ago by Hanson and Huxley (*19-21*), where they proposed that interdigitated thick and thin filament slide past each other during contraction and relaxation cycles (see Figure I-3B).

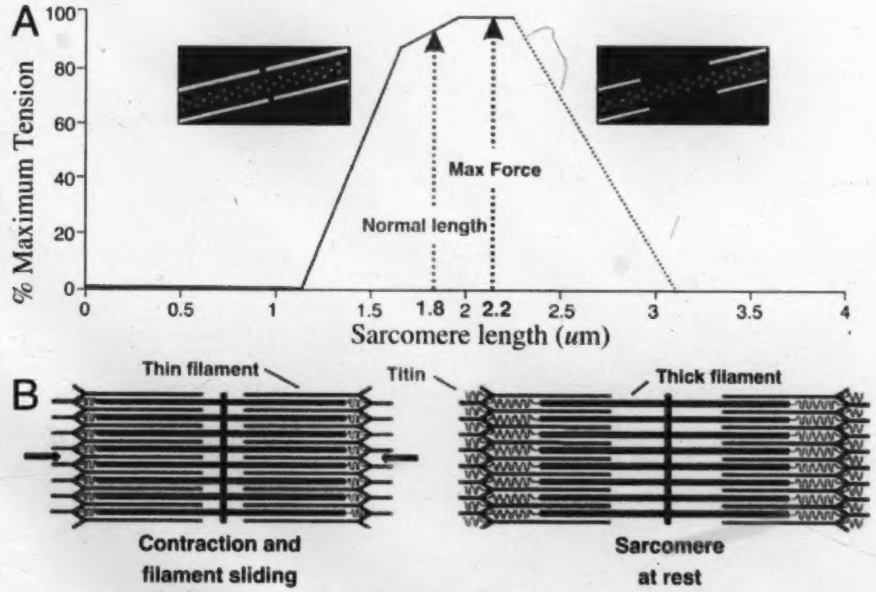

A
% Maximum Tension
100
80
60
40
20
0
Max Force
Normal length
0
0.5
1
1.5
1.8
2
2.2
2.5
3
3.5
4
Sarcomere length (*u*m)
B
Thin filament
Titin
Thick filament
Contraction and filament sliding
Sarcomere at rest

Figure I-3. **Maximal tension and the sliding filament hypothesis.** Diagram depicting (A) maximal force generation in the cardiac sarcomere and (B) the interpenetrating thick and thin filaments sliding past one another at rest (left) and during contraction (right). Figure adapted from (22).

The thick filament is composed almost entirely of the 220kDa heavy-chain protein *myosin*, which dimerizes via interactions of its C-terminal tail regions, into long α-helical coiled-coil domains of approximately 1500Å. The N-terminal region of each myosin forms an elongated head domain, to which two pairs of differing proteins are associated, namely the essential and regulatory light chains that vary in size from 15kDa-22kDa (see Figure 1-4). The long, fibrous, rod-like tails pack end to end in a staggered array of several hundred myosin molecules to form the thick filament, leaving the globular heads projecting to the sides on both ends. Each myosin head contains an ATPase domain that contains an ATP binding site, located within a V-shaped pocket on its surface. The central area of the thick filament where no myosin heads protrude out is referred to as the 'bare zone' (M-line) and can cover a distance of up to 1600Å, with myosin heads of opposite polarity positioned on either side.

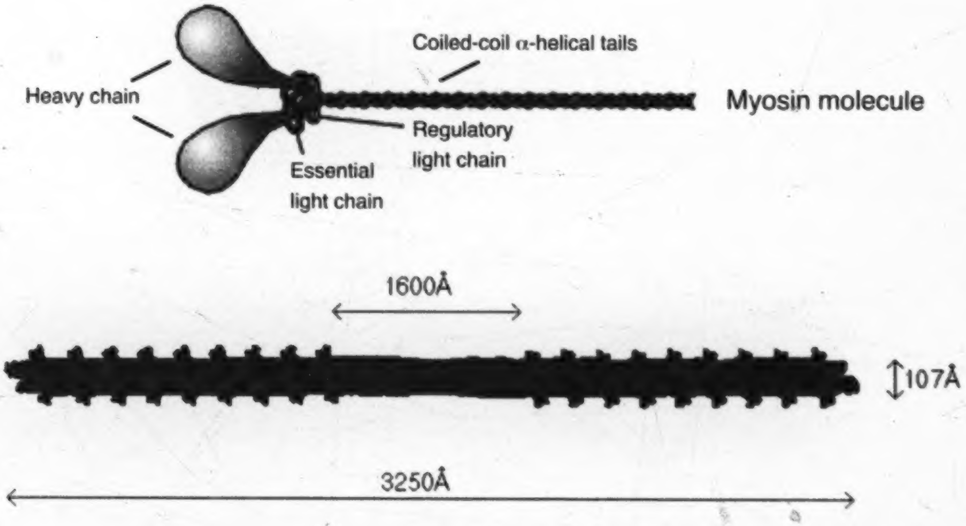


Figure 1-4. **Structure of the thick filament**. A single myosin cartoon molecule is shown at the top, complexed with regulatory and essential light chains. Interacting myosin subunits are shown at the bottom, in which several hundred myosin molecules form a staggered overlapping array (red) with their globular heads (blue) pointing away from the thick filament, with relative distances as indicated. Figure and legend adapted from (*17, 22*).

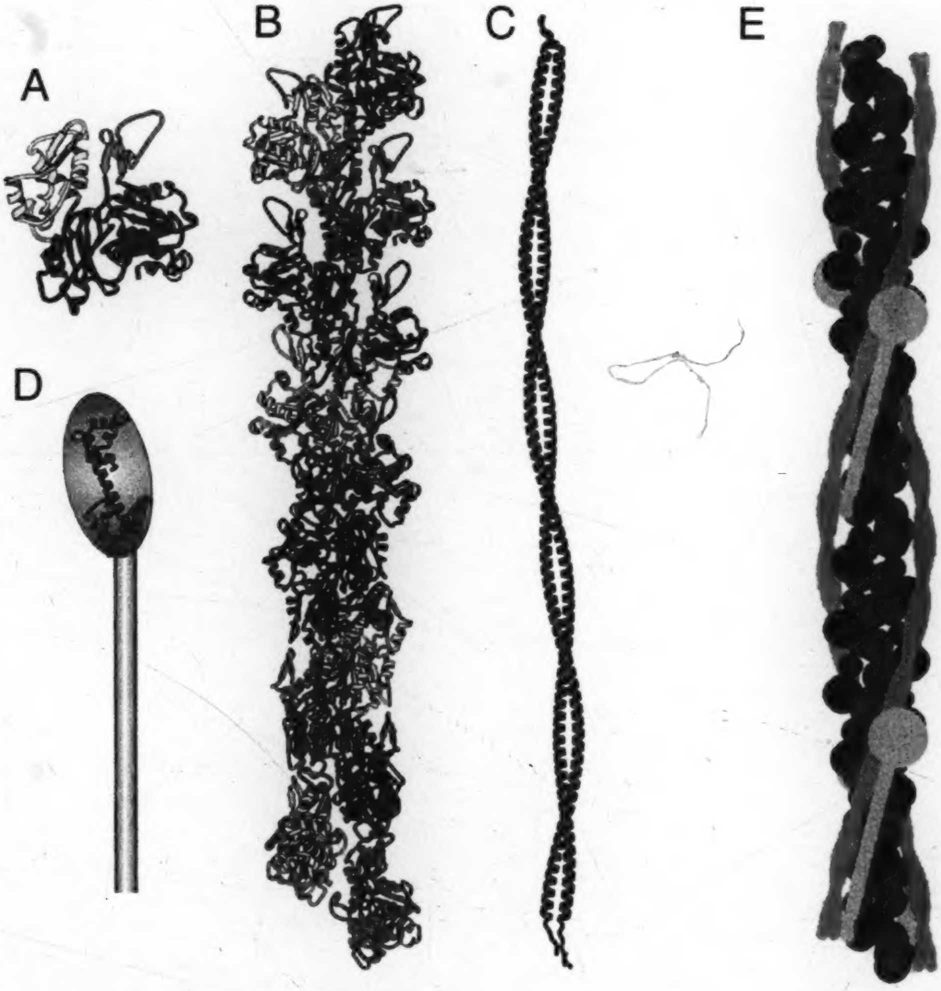


Figure 1-5. **Structure of the thin filament**. (A) Ribbon diagram of a single actin monomer with 4 specific subdomains (I-IV) colored as red, green, blue and yellow, respectively (*23*). (B) Ribbon diagram of the actin monomer placed into a 13/6 helix to model X-ray diffraction data of oriented actin filaments (*24*). (C) Ribbon diagram of the two-chain, coiled-coil α-helical structure of tropomyosin. (D) Illustration of the troponin complex, drawn to the same scale as panels A-C, where the known structure of troponin C (red) is overlaid on top of a blue oval which represents troponin I, attached to a green rod representing troponin T (*25*, *26*). (E) Cartoon diagram of the thin filament composed of subunits viewed in panels B-D. Actin is shown in red, tropomyosin strands are shown in green and the troponin complex is shown in blue. There is one tropomyosin and one troponin complex for every seven actin monomers within the actin filament. Figure and legend adapted from (*5*).

The thin filament is composed of three separate and distinct protein complexes, namely actin, tropomyosin and the troponin complex (see Figure I-5). Actin is a 375-residue globular protein, which polymerizes at physiological conditions to form long, fibrous filaments that constitute the core of the thin filament. The protein tropomyosin is a two-chain α-helical coiled-coils, which positions itself within the actin filament scaffold. Each tropomyosin dimer spans a distance of 400Å, with respective dimers associating in a head-to-tail overlapping fashion to form long, continuous cables that wind within the helical groove of the actin filament. Tropomyosin is positioned within the thin filament such that each tropomyosin molecule will come into contact with seven actin monomers. Each tropomyosin molecule binds a single heterotrimeric troponin complex. The troponin complex is comprised of three separate protein elements, namely troponin C which binds Ca^{2+} ions, troponin I which has a primary role in inhibiting muscle contraction, and troponin T which anchors the troponin complex onto tropomyosin.

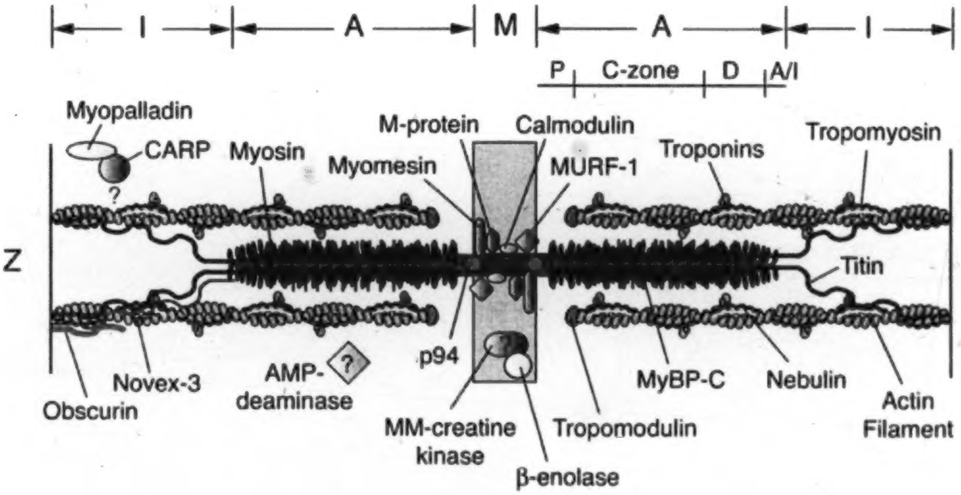

I
A
M
A
I
P
C-zone
D
A/I
Myopalladin
CARP
?
Myosin
M-protein
Myomesin
Calmodulin
MURF-1
Troponins
Tropomyosin
Z
Titin
Z
Novex-3
Obscurin
AMP-
?
deaminase
p94
MM-creatine
kinase
β-enolase
Tropomodulin
MyBP-C
Nebulin
Actin
Filament

Figure I-6. **Proteins of the sarcomere**. Molecular model representing the protein components of the I-band, A-band and M-line regions. Thin filaments composed of actin, tropomyosin and the troponin complex, span the I-band and interdigitate with the myosin ATPase heads of the thick filament. Other components know to associate with the sarcomere are as listed, with proteins whose binding sites are currently unknown shown with question marks. Figure and legend adapted from (7).

In addition to the thick and thin filament sarcomere proteins, there are additional protein complexes that play a role in muscle homeostasis (see Figure I-6). The myofibril is a dynamic environment containing many regulatory and structural proteins and enzymes including kinases, deaminases, scaffolding filament proteins (i.e. titin), calcium binding proteins (i.e. calmodulin), and others. In total, myosin and actin can account for up to 60% to 95% of total muscle protein by weight.

STRUCTURAL DIFFERENCES BETWEEN MUSCLE TISSUES

Cardiac, skeletal and smooth muscle tissues perform similar roles of contraction and relaxation cycles, however genetically, physiologically, functionally and structurally they are distinct. As the name implies, cardiac tissue is only expressed (isoform specific) within the heart, and cardiac cells are typically much smaller that those observed for skeletal muscle fibers. The transverse tubules (T-tubules) of cardiac muscle cells are short and broad, and encircle the sarcomeres at the Z-disk rather than at the zone of overlap (see Figure I-1). Cardiac muscle cells are almost totally dependent on aerobic respiration during contraction and therefore the sacroplasm contains large amounts of mitochondria and abundant reserves of myoglobin for oxygen storage. Cardiac muscle cells also contain specialized sites known as *intercalating discs* (see Figure I-6), which are present on the cell membranes between neighboring muscle cells that are extensively intertwined, functioning to maintain the three dimensional structure of the tissue. The intercalating discs, along with gap junctions and desmosomes, are arranged as to create a direct electrical connection, allowing action potentials to move quickly from one cardiocyte to the next. Because the myofibrils are attached to the cell membrane at the site of the intercalating discs, this allows cardiac muscle cells to contract together for maximal efficiency, thus the entire cardiac muscle tissue resembles a single, enormous muscle cell. Cardiac tissue contracts without neural stimulation, with timing of contractions determined by specialized cells referred to as pacemaker cells.

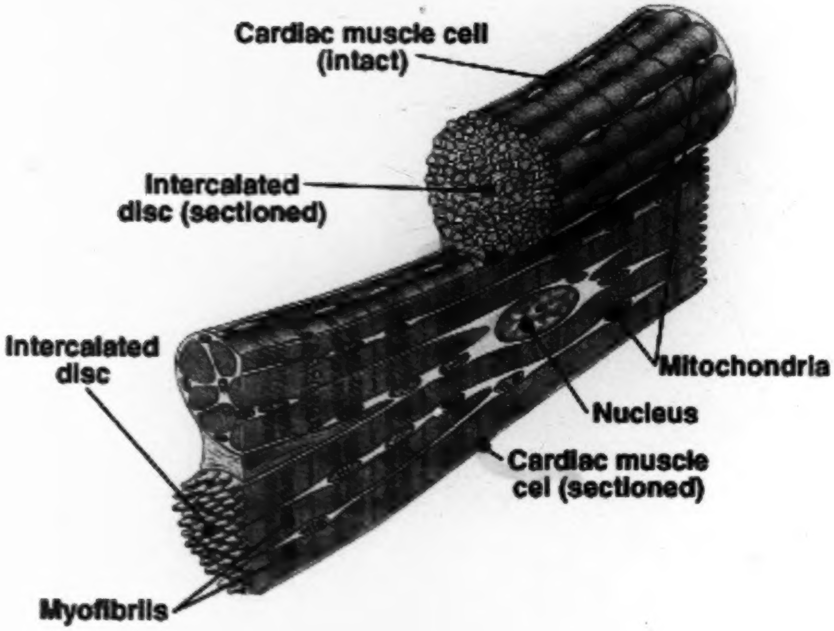

Cardiac muscle cell
(intact)
Intercalated
disc (sectioned)
Intercalated
disc
Mitochondria
Nucleus
Cardiac muscle
cel (sectioned)
Myofibrils

Figure I-7. **Intercalating discs**. Unique to cardiomyocytes, intercalating discs are found on cell membranes between two adjacent cardiac muscle cells and help to propagate functionality of adjacent cells. Figure adapted from (*22*).

THE MECHANISM OF MUSCLE CONTRACTION

Force generating molecular interactions originate when a signal from pacemaker cells initiates an action potential into a series of T-tubules, which extend inward from the plasma membrane around the myofibrils. This signal is then relayed to the sarcoplasmic reticulum (SR), an organelle that contains high concentrations of cellular calcium (*27, 28*). Within the junctional region, large Ca^{2+} release channels (CaRC) extend outwards like pillars from the SR, making contacts with the T-tubules (*29, 30*). When voltage sensitive proteins within the T-tubules sense an action potential, they trigger the CaRCs to open, allowing Ca^{2+} ions to escape from the SR, which flood onto the myofibril filaments to initiate contraction (see Figure I-8).

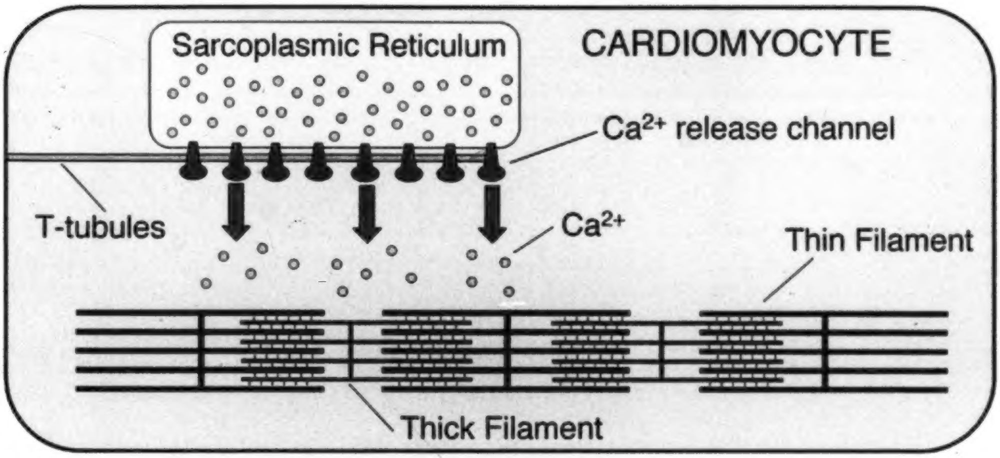


Figure I-8. **Calcium release from the sarcoplasmic reticulum.** Cartoon diagram of a simplified cardiomyocyte depicting the sudden rise of calcium concentrations released from the sarcoplasmic reticulum via Ca^{2+} release channels, initiating muscular contraction.

The calcium dependence of cardiac contraction is due entirely to the accessory protein, the troponin complex, found associated within the thin filament (*31*). The binding of calcium to troponin C initiates domain conformational changes, which are propagated via the troponin complex throughout the entire thin filament (*32*). The thin filament undergoes large structural changes in which tropomyosin is repositioned from the inner edge of the outer domain of actin to a position on the outer edge of the inner domain of actin, essentially moving deeper into the helical cleft of the filamentous actin scaffolding (*13, 33*) (see Figure I-9). Each actin monomer on the thin filament is capable of binding a myosin head on the thick filament via ionic interactions and the association of hydrophobic patches on the exposed surface of each subunit. The tropomyosin-troponin complex regulates muscle contraction by controlling the availability of the actin binding sites with their binding partner myosin subunits. Together, tropomyosin and troponin I sit on top of the actin-myosin binding site on the thin filament and inhibit muscular contraction in the absence of calcium signaling through a steric-blocking mechanism.

Following the relocation of tropomyosin-troponin off the actin-myosin binding site, ATP-dependent contraction is capable of proceeding via cross-bridge interactions between the thick filament myosin heads and the thin filament actin subunits (see Figure I-10). The sliding filament hypothesis is proven correct via repeating cyclic interactions of cross-bridge formation and release between the myosin and actin domains (*20, 21*). ATP binds to a myosin head that causes the actin-binding site to open and release a bound actin. The active site then closes around the ATP, resulting in hydrolysis that 'cocks' the myosin head into a high-energy conformation. The myosin head binds weakly to a new actin monomer, which is closer to the Z-disk than the previous one it had been attached to. Myosin then releases an inorganic phosphate molecule (P_i) resulting in a conformational shift that increases its affinity for actin. The power stroke by the myosin head follows, where the heads C-terminal tail is swept upwards of 60Å towards the Z-disk relative to the actin-binding site on the head. ADP is then released by the myosin head, allowing the contraction cycle to repeat itself over and over. The many hundreds of myosin heads present on each thick filament can individually perform this function numerous times per second. Contraction cycles continue to occur until a desired amount of force has been achieved, followed by cardiac relaxation initiated by a loss of the calcium signal via a re-uptake of calcium ions back into the SR. It is through the 'sliding' of the two filaments across one another that myosin has been referred to as having the ability to 'walk' or 'row' along adjacent thin filaments towards the Z-disk.

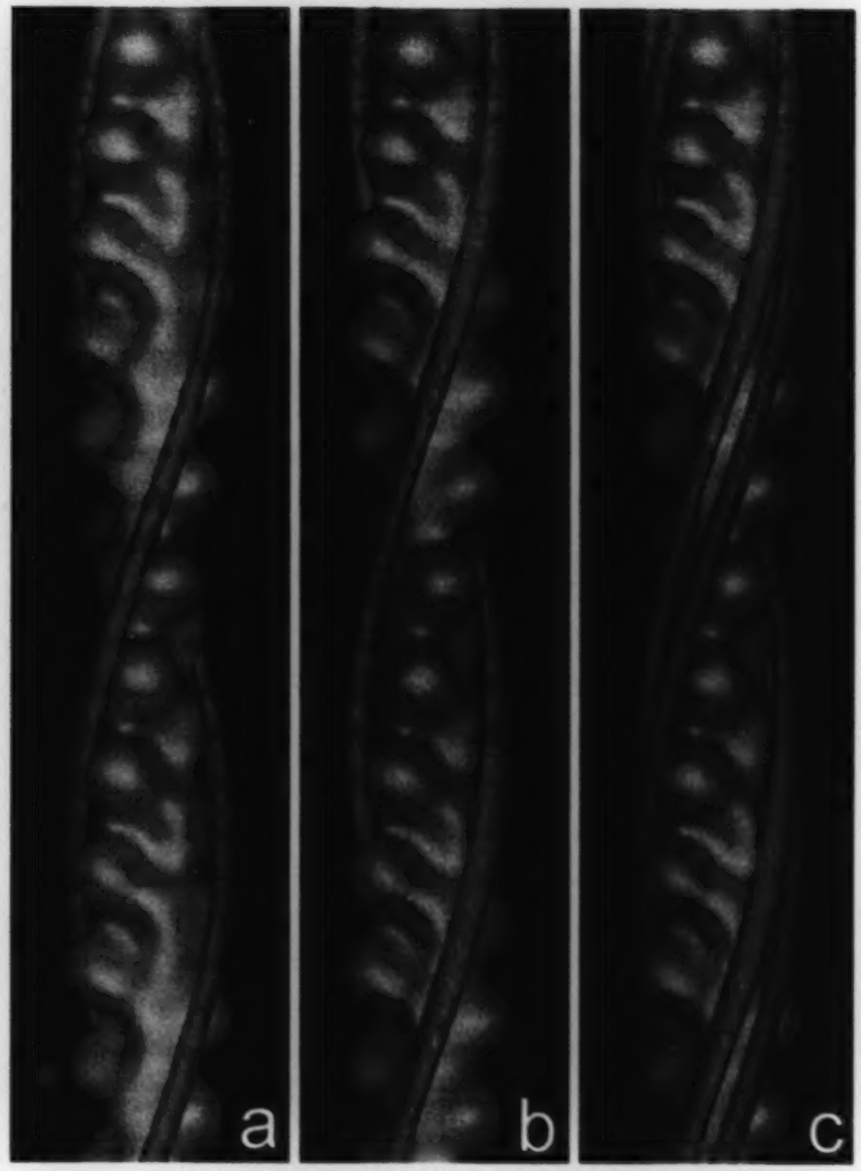


Figure 1-9. **Ca^{2+}-induced thin filament repositioning of tropomyosin.** Surface rendered reconstructions of the thin filament showing the positioning of tropomyosin in the presence of (a) EGTA and (b) Ca^{2+}. Panel (c) displays a superimposition of the two states, demonstrating the repositioning of tropomyosin on the thin filament from a position on the inner edge of the outer domain of actin in the apo-state (red), to a position along the outer edge of the inner domain of actin in the Ca^{2+}-activated state (green). Filamentous actin is shown in gold. Figure and legend adapted from (*13, 33*).

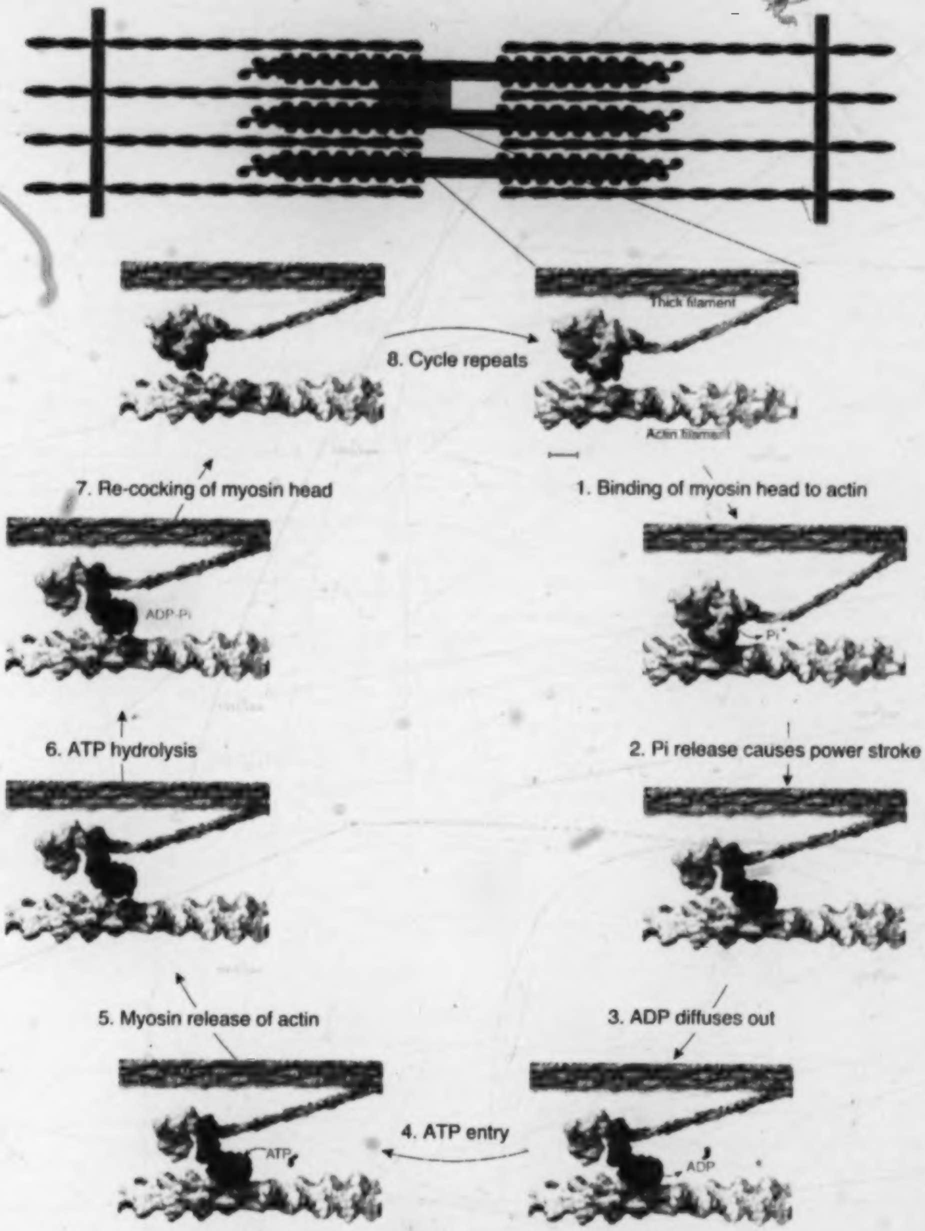

Thick filament
Actin filament
8. Cycle repeats
1. Binding of myosin head to actin
Pi
2. Pi release causes power stroke
3. ADP diffuses out
ADP
4. ATP entry
ATP
5. Myosin release of actin
6. ATP hydrolysis
ADP·Pi
7. Re-cocking of myosin head

Figure 1-10. **Mechanism of force generation in muscle.** 8-step cycle of contraction where myosin heads ‘walk’ up the actin subunits (green) of the thin filament in a unidirectional cyclic process that is driven by the hydrolysis of ATP, causing structural changes within the myosin heads (shown in red) to generate the power stroke via cross-bridge formation. Figure and legend were adapted from (*17, 34*).

TROPONIN C – CALCIUM CONTROL OF MUSCULAR CONTRACTION

Troponin C is the Ca^{2+}-binding component of the troponin complex. It is the best structurally characterized component of troponin with data from various species of bird, mammals and invertebrates having been described in the literature since 1985 (*25, 26, 35, 36*). Cardiac troponin C is a 161-residue protein of molecular weight 18.4kDa, which contains an iso-electric point of 4.2, and shares sequence and structural homology to other troponin C isoforms (*37, 38*). cTnC contains four helix-loop-helix structural motifs referred to as EF-hand domains (*35*), each consisting of a 12-residue Ca^{2+}-binding loop flanked on either side by α-helices (*39-41*). Ca^{2+}-chelation occurs via coordination with carboxyl and/or hydroxyl groups of side-chain residues and carbonyl backbone amide groups, with 6 of the 12 residues participating in complexation (residues 1, 3, 5, 7, 9, 12). Ca^{2+} affinity and specificity is determined by both the primary structure of the loop and the composition/interactions of the flanking helices within individual EF-hands (*38, 42-44*).

cTnC is a dumbbell-like molecule consisting of two globular domains: the N-lobe and C-lobe, which are connected by a central, flexible linker region (Figure I-11). Both the N-lobe and C-lobe each contain two EF-hand motifs respectively (*45*). The N-lobe contains two low-affinity Ca^{2+}-binding loop (Sites I and II), however site I contains an insertion and non-conservative mutations rendering it unable to bind Ca^{2+}. The C-lobe contains two high-affinity Ca^{2+}-binding loops (Sites III and IV). The binding affinity of site II within the N-lobe ($K_a \approx 10^6$ M^{-1}) is significantly lower than sites III and IV within the C-lobe ($K_a \approx 10^7$ M^{-1}), although the N-lobe possesses higher specificity (*46*). Unlike the C-lobe that binds both Ca^{2+} and Mg^{2+} ($K_a \approx 10^3$ M^{-1}), the N-lobe does not bind Mg^{2+} and is therefore designated as a Ca^{2+}-specific low affinity Ca^{2+}-binding site (*47-49*). The N-lobe contains 90-residues consisting of five α-helices where helices A/B flank the defunct Ca^{2+}-binding site I, and helices C/D flank the low-affinity Ca^{2+}-binding site II. In addition the N-lobe contains a helical extension at the N-terminus of the protein, the N-helix, which plays an important role in the proper alignment of the N-lobe with cTnC and calcium binding. The C-lobe contains 71 residues consisting of four α-helices, where helices E/F and G/H flank the Ca^{2+}-binding sites III and IV respectively (see Figure I-11). The N-lobe is structured in the absence of Ca^{2+} unlike the C-lobe, which requires divalent cations Ca^{2+} and/or Mg^{2+} for tertiary structure formation. The structure of cTnC has been solved by NMR, which reveals the flexible nature of the D/E linker region (*50*).

```
        N-helix       A-helix                 B-helix         C-helix
  1 AADIYKAAVE QLTEEQKNEF KAAFDIFVLG AEDGSISTKE LGKVMRMLGQ NPTPEELQEM 60

                    D-helix              E-helix             F-helix
 61 IDEVDEDGSG TVDFDEFLVM MVRSMKDDSK GKTEEELSDL FRMFDKNADG YIDLEELKIM 121

               G-helix         H-helix
121 LQATGETITE DDIEELMKDG DKNNDGRIDY DEFLEFMKGV E                     161
```

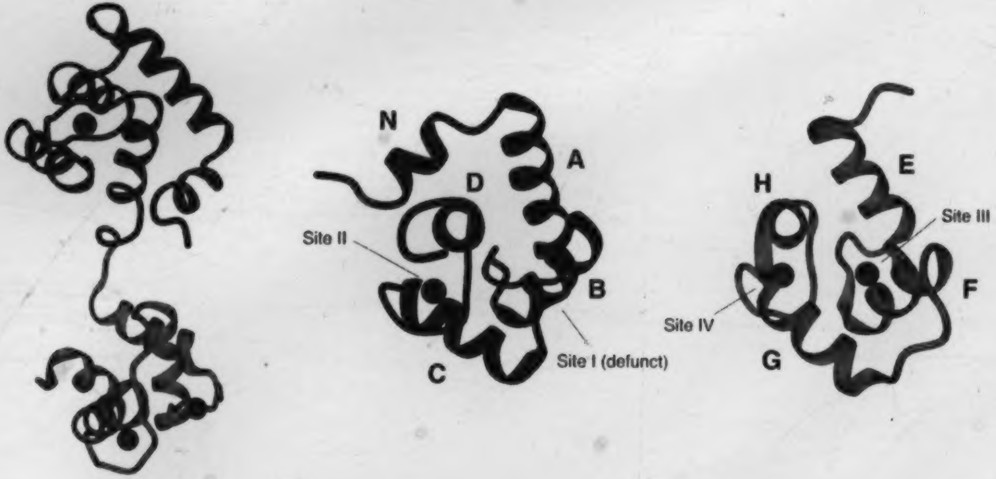


Figure I-11. **Structure of cardiac troponin C**. Primary amino acid sequence of cardiac TnC [top]. Ribbon backbone diagram of the NMR solution structure of calcium-saturated cardiac TnC (1AJ4) where the N-lobe is colored in blue, the C-lobe in green and bound Ca^{2+}-ions in red [left]. Rotation of the individual N- and C-lobe with indicated helices and positioning of Ca^{2+}-binding sites [right]. Figure adapted from (*50*).

During cardiac relaxation, there is a low concentration of Ca^{2+} within the myofibril and only the C-lobe sites are saturated with Mg^{2+} and/or residual Ca^{2+}-ions. Due to the inability of the C-lobe to selectively distinguish between divalent cations, it cannot directly participate in the regulation of cardiac contraction. These sites delegate more of a structural role for the C-lobe in providing fixation of cTnC to other components of the thin filament (*51, 52*).

Under physiological resting conditions (i.e. N-lobe apo, C-lobe saturated), the structures of the N-lobe and C-lobe differ significantly as the N-lobe is in a 'closed' state, whereas the C-lobe is in a constituently 'open' state. A closed state is defined by the interhelical angle between the A/B and C/D helices of the N-lobe, in that the calcium binding sites can be thought of as a hinge domain, whereby the B and C helices can swing out into an open stated, revealing hydrophobic surface area for binding (*37*).

Cardiac contraction is initiated by an SR response to increase sarcoplasmic calcium concentrations. Calcium ions flood into the myofibril, allowing calcium ions to intercalate within the thin filament. The N-lobe is the calcium switch, as only during the increased calcium concentrations will its low-affinity Ca^{2+}-binding site II bind to a single Ca^{2+} ion. The structural consequences of calcium binding to site II are minimal, with the N-lobe remaining in the closed state (see Figure I-12 and Table I-1) (*50*). Although closed in the calcium-saturated state, the affinity of the N-lobe for a section of troponin I called the 'switch region' ($cTnI_{147-163}$) is markedly increased (*53*) and allows for a binding event to occur via hydrophobic interactions. It is this binding event that alone is sufficient to initiate the structural changes to alleviate the inhibition of muscle contraction.

It is due to this ability of cTnC to sense calcium levels that it has been referred to as the 'calcium trigger', whereby it is literally the smoking gun of muscular contraction. Cardiac TnC is therefore the quarterback of the team, dictating its will on all the other components of cardiocyte contraction.

Table I-1.
Interhelical angles of N-lobe

N-lobe state[b]	Interhelical angles (°)[a] A-B	C-D	Accession code
N-lobe • apo	136 ± 3	129 ± 5	1SPY
N-lobe • Ca^{2+}	132 ± 3	117 ± 3	1AP4
N-lobe • Ca^{2+} • $cTnI_{147-163}$	102 ± 5	96 ± 5	1MXL

[a] A large angle defines a closed conformation, whereas a small angle defines an open conformation. The axis for an α-helix is defined by two points, taken as the average coordinates of the first and last 11 backbone atoms of the α-helix.

[b] Structures taken from solved NMR solution structures of the N-lobe.

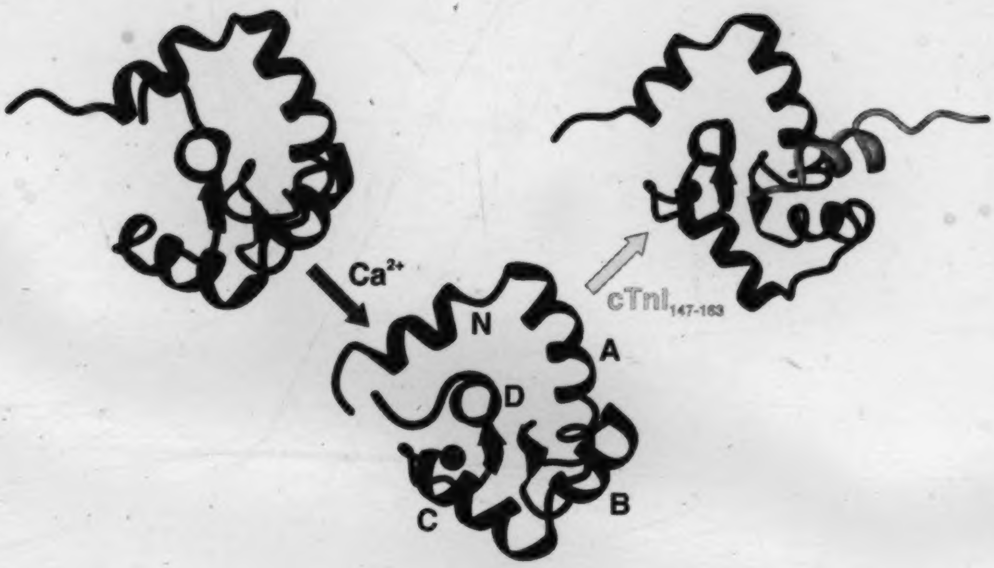


Figure I-12. **Calcium binding to N-lobe**. Ribbon backbone diagram depicting the step-wise structural consequences of the N-lobe (blue) to Ca^{2+}-binding. Upon the binding of a single Ca^{2+}-ion (red) into site II, the N-lobe remains in the 'closed' calcium-saturated state, however binding affinity to the 'switch region' of troponin I (yellow - $cTnI_{147-163}$) is increased and an 'open' state of the N-lobe is revealed upon association of the two domains via hydrophobic contacts. Figure adapted from (*53, 54*).

TROPONIN T –TROPOMYOSIN BINDING AND CALCIUM SENSITIVITY

Cardiac troponin T is the largest and longest component of the troponin complex and possesses extensive and diverse interactions with its thin filament neighbors including tropomyosin, actin and its troponin brethren: cTnC and cTnI (*6, 55-57*). cTnT can be separated into two domains: T1 (residues 1-183) and T2 (residues 184-290) (*58*). T1 forms the N-terminal tail region of cTnT, binding to the tropomyosin-tropomyosin overlap region and conferring cooperative myosin binding to the thin filament, and also possesses a role in myofilament activation and relaxation (*59-61*). The N-terminal region of T1 appears to be important in maintaining tropomyosin in the blocked state with steric block of the actin cross-bridge reaction. T2 participates in interactions with cTnC-cTnI in sensing Ca^{2+} concentrations and inhibiting force development (*59*). cTnT is an approximately 290-residue protein that has variable isoforms ($cTnT_1$-$cTnT_4$), which arise from alternative splicing of TnT transcripts during human heart development and in heart failure (*62*). It is the least structurally understood component of the troponin complex, with little structural data regarding the T1 domain (*51*).

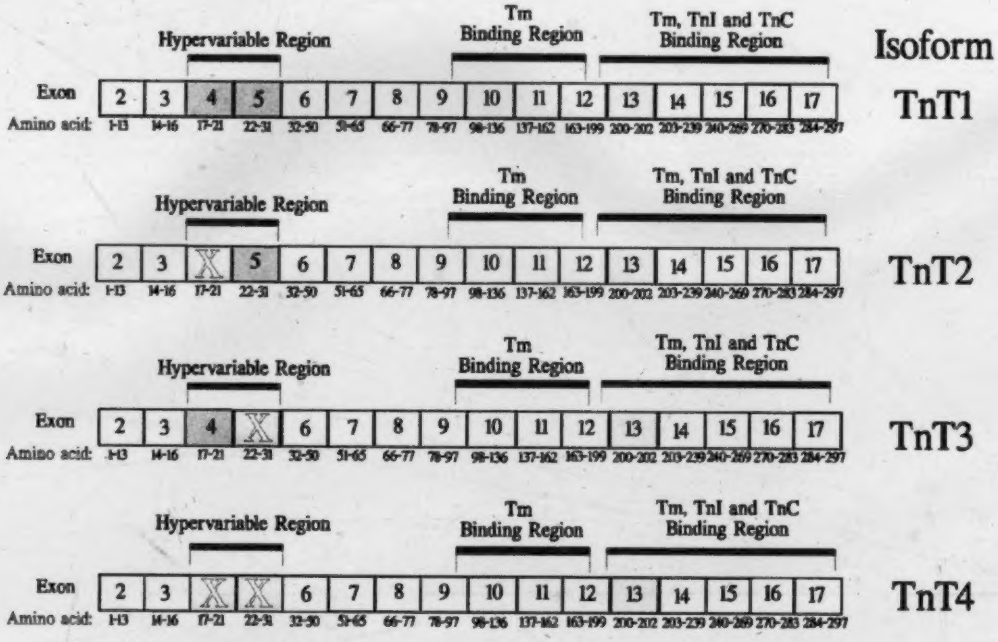
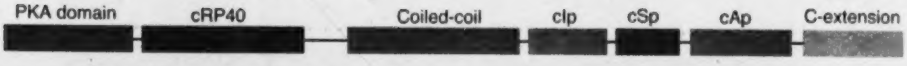

Isoform
Hypervariable Region
Tm Binding Region
Tm, TnI and TnC Binding Region
Exon
Amino acid:
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
1-13
14-16
17-21
22-31
32-50
51-65
66-77
78-97
98-136
137-162
163-199
200-202
203-239
240-269
270-283
284-297
TnT1
TnT2
TnT3
TnT4

Figure I-13. **Cardiac troponin T variability**. Schematic diagram of isoform splicing, with areas of deletion marked with a large X. Figure taken from (*61*).

TROPONIN I – INHIBITION OF CARDIAC CONTRACTION

Cardiac troponin I is a 24kDa, 210-residue basic protein with a high iso-electric point of 9.8, and was originally named for its ability to inhibit actin-activated myosin ATPase activity (Figure I-14). It is located on chromosome 19 and consists of 8 exons (*63, 64*). There are seven distinct functional regions within cTnI (N- to C-terminal): (1) a PKA-dependent phosphorylation N-terminal extension domain; (2) a 37-residue spanning region (cRP40) that binds to the C-lobe of cTnC; (3) a region that interacts with the T2-section of cTnT via a coiled-coil motif; (4) a basic inhibitory domain (cIp) that is the first actin-binding region; (5) a regulatory or triggering 'switch' region (cSp); (6) a second actin-binding domain (cAp); and (7) a C-terminal extension region.


PKA domain
cRP40
Coiled-coil
cIp
cSp
cAp
C-extension

```
  1 MADGSSDAAR EPRPAPAPIR RRSSNYRAYA TEPHAKKKSK ISASRKLQLK  50
 51 TLLLQIAKQE LEREAEERRG EKGRALSTRC QPLELAGLGF AELQDLCRQL 100
101 HARVDKVDEE RYDIEAKVTK NITEIADLTQ KIFDLRGKFK RPTLRRVRIS 150
151 ADAMMQALLG ARAKESLDLR AHLKQVKKED TEKENREVGD WRKNIDALSG 200
201 MEGRKKKFES                                             210
```

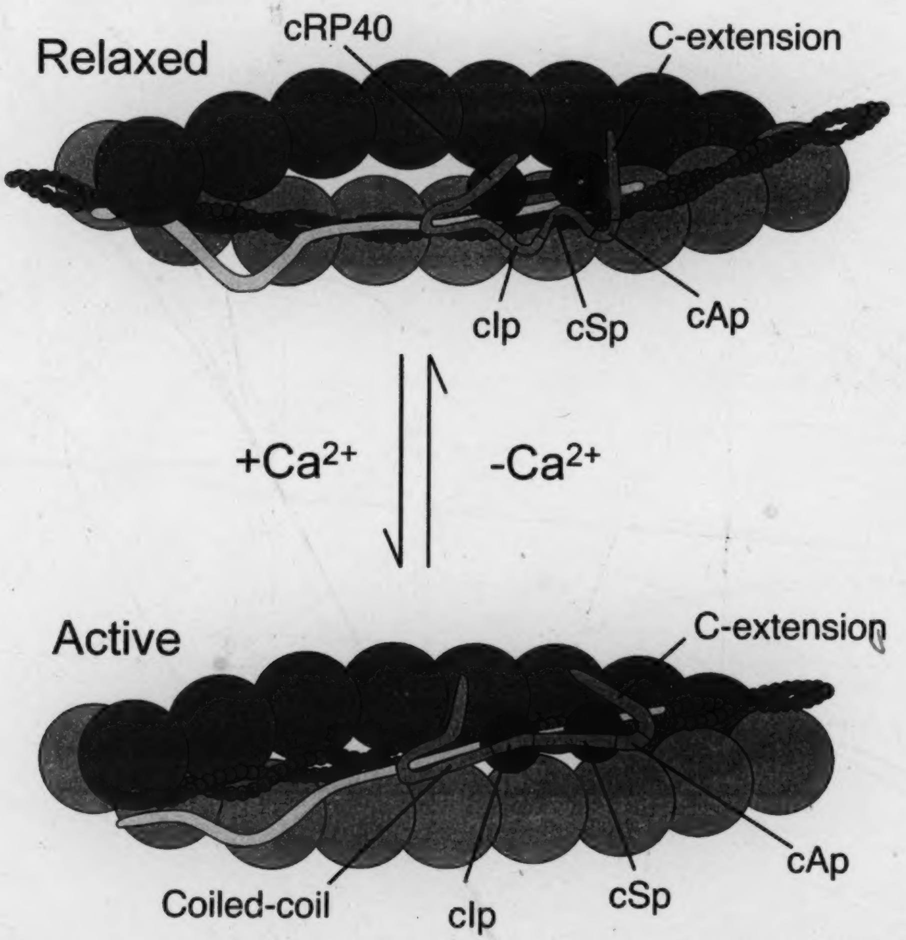

Relaxed
cRP40
C-extension
cIp
cSp
cAp
+Ca^{2+}
-Ca^{2+}
Active
C-extension
Coiled-coil
cIp
cSp
cAp

Figure I-14. **Calcium induced structural changes of the troponin complex.** Primary amino acid sequence of the 7 distinct domains of cTnI (above). Tripett cartoon model depicting tropomyosin-troponin I domain re-arrangements upon Ca^{2+}-induced cardiac contraction (below). Troponin I regions (cRP40, cIp, cSp, cAp and C-terminal extension) are as indicated. Actin is drawn in green, cTnT in yellow, tropomyosin in grey, cTnC in red and cTnI in blue. Figure adapted from (*65, 66*).

The main function of cTnI is to inhibit myosin ATPase activity at low cytoplasmic [Ca^{2+}] via interactions with actin (*6, 13, 67*). Troponin I effects are so dominant that even in the absence of tropomyosin and other troponin components *in vitro*; inhibition is still observed (*68, 69*). Maximal inhibition is achieved in the presence of tropomyosin at an actin:TnI ratio of 7:1. The periodic distribution of cTnI along the thin filament indicates that tropomyosin amplifies the inhibitory mechanism, as well serves to direct cTnI placement to an optimal thin filament location within the sarcomere (*9, 10*). The inhibitory region is the critical actin-binding region of cTnI (*68, 70*), whose sequence is highly conserved throughout all muscle types. cTnI undergoes substantial changes upon the initiation of muscular contraction: (1) cSp binding to the N-lobe of cTnC (see figure I-12); (2) the cIp region moving away from the actin-binding site; (3) the cAp region moving away from the second actin-binding site; and (4) movement with tropomyosin-troponin complex into the helical groove of the actin filament (see figure I-14).

During embryogenesis and development, both the cardiac and the slow skeletal isoforms of TnI are expressed in the heart of the human fetus. Following birth, the expression of the slow skeletal isoform is repressed, whereas the cardiac isoform is enhanced. By the ninth month of development (birth), the cardiac isoform is the exclusively expressed isoform (*71*). The cardiac isoform remains the only expressed isoform in cardiac tissues for the entire lifetime of an individual.

If cTnC is the quarterback of muscular contraction then cTnI is the receiver: accepting the signal from cTnC and marching a pre-designed play forward to a touchdown resulting in a contraction event. The 'I' in this team is arguably the most important and integral component of the cardiac contraction cycle, as it scores a touchdown every time it receives the ball. Due to its high fidelity of pass completion, cTnI has garnered a large amount of structural studies over the years in attempts to elucidate physiological and structural events at molecular resolution. In addition to early biochemical work by various groups in defining the seven-individual interacting sub-regions of cTnI (as defined above), many groups have recently attempted to define structural data via a combination of crystallography, NMR and fluorescent techniques. The troponin complex is a highly flexible entity that has limited solubility in low ionic strength solutions, which has limited its ability to produce high-quality crystals for x-ray structural analysis. To alleviate this obstacle, structural work has been primarily limited to domain interactions complexed with smaller domain fragments of parts or in combinations of the 7-sub-regions of cTnI and/or cTnT, by NMR and crystallography (*47, 48, 53, 65, 72-88*).

THE INHIBITORY REGION OF TROPONIN I

The inhibitory region is a critical domain for cardiac rhythm as it is the sole region of cTnI that possesses the ability to inhibit cardiac contraction. Originally isolated by Head and Perry using proteolytic and CNBr cleavage techniques (*31*), and subsequently defined by Talbot (*68*), the inhibitory region has been mapped to a central region of the cTnI sequence, and corresponds to residues Thr128-Arg147. The inhibitory region alone contains the minimum sequence required to fully inhibit myosin ATPase activity on the thin filament during the contraction cycle (*70*), thus interactions of the inhibitory region are of key importance in a complete understanding of the regulation of contraction. During cardiac relaxation, the inhibitory region, along with tropomyosin, interacts with actin on the thin filament, blocking the binding site for myosin ATPase (*89-91*). It is this blocking event that is sufficient to completely inhibit the ability of myosin to perform ATP-dependent cardiac contraction (*68, 69*). Upon an intracellular influx of calcium, contraction is initiated as the inhibitory region relocates from actin to cTnC, signaling tropomyosin movement into the actin-helical groove (*13, 33*). The actin binding sites for myosin are exposed resulting in the ability of the thick filament to slide over the thin filament via cross-bridge cycles (*17, 34*).

The amino acid sequence and physiological effects of the inhibitory region have been known for many years. Elucidation of associated proteins and binding sites have been mapped through a series of biophysical experimentation such as cross-linking, NMR, fluorescence and mutational analysis (*65, 92*). From a solely structural standpoint, more is known regarding the inhibitory region in the contracted state, as *in vitro* experimentation has proven to be monumentally easier given the excessive solubility of the inhibitory region and the globular nature of cTnC. Investigations of the inhibitory region bound to actin in the relaxed state has proven to be difficult due to the preference of F-actin to polymerize at physiological conditions, limiting most studies solely to *in vivo* experimentation. The inhibitory region has been shown to interact with the C-lobe and the central D/E linker region of cTnC (*73, 93*), however no high-resolution structural data is available regarding the atomic placement of the inhibitory region in the contracted state. Since the inhibitory region contains a large percentage of basic residues (8 out of 20) and its cTnC binding partner contains a large number of acidic residues (C-lobe: 32 out of 71), electrostatic interactions may play a role in the interaction (*85*).

The ability of other cTnI domains to associate with cTnC has yielded valuable structural information in explaining muscular contraction at atomic resolution. The structure of the cSp region is known to bind to the N-lobe in the calcium-saturated site (see Figure I-12), and the cRP40 region is known to bind in the hydrophobic cleft of the C-lobe (*53, 72, 74, 84*). Unfortunately, high-resolution data regarding the placement of the inhibitory region on cTnC during the contraction cycle is unavailable. The elucidation of this structure would be a large step forward in obtaining a greater understanding of the contractility machinery at the atomic level, however the only data currently available regarding the structural orientation of this domain stems from indirect measurements such as fluorescence and various proposed models (*72, 94-99*). For the structural elucidation of these cTnI domains, a critical assumption is made in that an isolated domain behaves identically to how it would in the full complex.

TROPONIN I – MUTATION AND CARDIOMYOPATHY

Cardiomyopathy is a serious disease in which cardiac muscle becomes inflamed and does not possess the ability to function at a physiological level. There are three main types of cardiomyopathy: dilated, hypertrophic and restrictive (*60, 100*). Dilated is the most common form, where the heart cavity is enlarged and stretched, resulting in abnormal pumping and altered cardiac output. Restrictive is the least common form and is characterized by a myocardium (heart muscle) that has ventricles that have become excessively rigid, rendering if more difficult for ventricles to properly fill with blood between heartbeats. Hypertrophic cardiomyopathy (HC) is characterized by the enlarging of the left ventricular muscle mass into 'hypertrophies', which vastly affects cardiac output and function. In the majority of the cases, HC is hereditary and is most prevalent in young adults.

cTnI disarray has been shown to be one of the factors in leading to a HC phenotype. For example, familial hypertrophic cardiomyopathy (FHC), one of the most frequently occurring inherited cardiac disorders, which affects up to 1 in 500 of the population, has been identified to be a sarcomere disease (for reviews, see (*60, 100*)). It is characterized by left ventricular hypertrophy, myofibril disarray, sudden cardiac death, and is believed to be caused by mutations in certain contractile protein genes, including cTnI. To date, six missense mutations (R144G, R144Q, R161W, S198N, G202S, and K205Q) on cTnI have been reported along with mutations causing the deletion of one codon (ΔK183) and a deletion of Exon8 (Δexon8) (*100*). Among those, R144G and R144Q are located in the key inhibitory region of cTnI. A truncation of cTnI has been associated with myocardial stunning (*101, 102*) and measuring the serum level of cTnI has become the standard of care in the diagnosis of myocardial injury (*89, 103-105*). Not surprisingly, most mutations that directly affect cTnI activity all involve a non-conserved mutation where there is a loss of a charged residue. The consequences involving altered charged interactions with the two acidic cTnC domains may be extremely physiologically hazardous to the long-term efficiency of cardiac health and viability.

TROPONIN I – POST-TRANSLATIONAL MODIFICATION

Covalent modification of cTnI has been shown to dramatically affect cardiac performance resulting in markedly variable responses to relaxation rate, calcium sensitivity and/or force generation. Phosphorylation of the thin filament regulatory proteins is maintained and controlled predominantly by protein kinase A (PKA) and protein kinase C (PKC), of which cTnI is a known target.

PKA is a known effector of β-andrenergic stimulation within cardiocytes, where there is a rapid formation of the intracellular secondary messenger cyclic AMP (cAMP) which causes the direct activation of PKA via an auto-phosphorylation mechanism. PKA has two key phosphorylation sites within cTnI (Ser22 and Ser23), which are important regulatory residues in response to stimulation. β-andrenergic stimulation of cardiocytes results in contractile performance that is markedly enhanced (*106-108*), which can be lead to increases in cardiac pressure generation, relaxation rates and cardiac contraction frequency. For example, in a resting human heart the average heart rate is 70 to 80 beats per minute (bpm) with left ventricular cardiac output of 4 to 5 L/min, however during β-andrenergic stimulation heart rates can exceed 220 bpm with an output increase to greater than 30 L/min (*106*). This ability of the human heart to alter cardiac function by a broad range is of central physiological significance to overall biological fitness to meet the changing metabolic demands from the resting state to a maximal 'flight or fight' stress response state (*109*). To accomplish the above mentioned responses to β-andrenergic stimulation, the myocardium must be able to relax from a higher force over a shorter time interval in order to preserve proper diastolic function of the heart ventricles (*109*).

In addition to β-andrenergic stimulation of cTnI, an alternate regulation pathway of cTnI activity is available via phosphorylation of residues Ser41, Ser43 and Thr142 by PKC (*110, 111*). With the earlier stated hypotheses of electrostatic interactions playing a potential role in the interaction of the inhibitory region with cTnC, perhaps it is not surprising that a phosphorylation event occurs within this region (Thr142) as this may make a large negative electrostatic contribution to the overall binding. It is through the fine-tuning of the phosphorylation/dephosphorylation pathway(s) that cardiac tissues are able to directly respond to environmental stresses via regulation in both the long- and short-term. The effect of cTnI phosphorylation is primarily a downward modulation of cardiac contractility, either by decreased actomyosin ATPase activity, which is mediated by PKC, or by increased binding of cTnI to the thin filament, mediated by PKA.

TROPONIN I – STRUCTURAL INSIGHTS FROM A CRYSTAL STRUCTURE

In the July 2003 issue of Nature, the Maeda group succeeded where many other have failed in the past and made an contribution to the field of cardiac structural research by producing the first high-resolution crystal structure of a 46kDa and a 52kDa 'core' region of the troponin complex, which included cTnC, the T2 domain of cTnT and regions of cTnI (*51*). Within this crystal structure, domain orientations of the troponin complex were first visualized in the calcium-saturated state (see figure I-15A). The crystal structure correlated well with previous structural elucidations of smaller components of cTnI in complex with various troponin fragments by NMR and crystallography, confirming earlier assumptions that the previously reported subunit complexes involving isolated domains retained all biologically relevant structural data of the whole complex.

The Maeda structures (1J1D.pdb & 1J1E.pdb) were instrumental in showing the coiled-coil interactions between cTnT-T2 and cTnI, and is referred to as the IT arm. The specifics of this crystal structure are outside the realm of this thesis introduction, and will be addressed in detail in subsequent chapters and discussions, however based upon this structure a model can be proposed for thin filament mediated cardiac contraction (see Figure I-15C). It is important to note that the crystal structure failed to produce electron density of the critical inhibitory region, presumably due to flexibility.

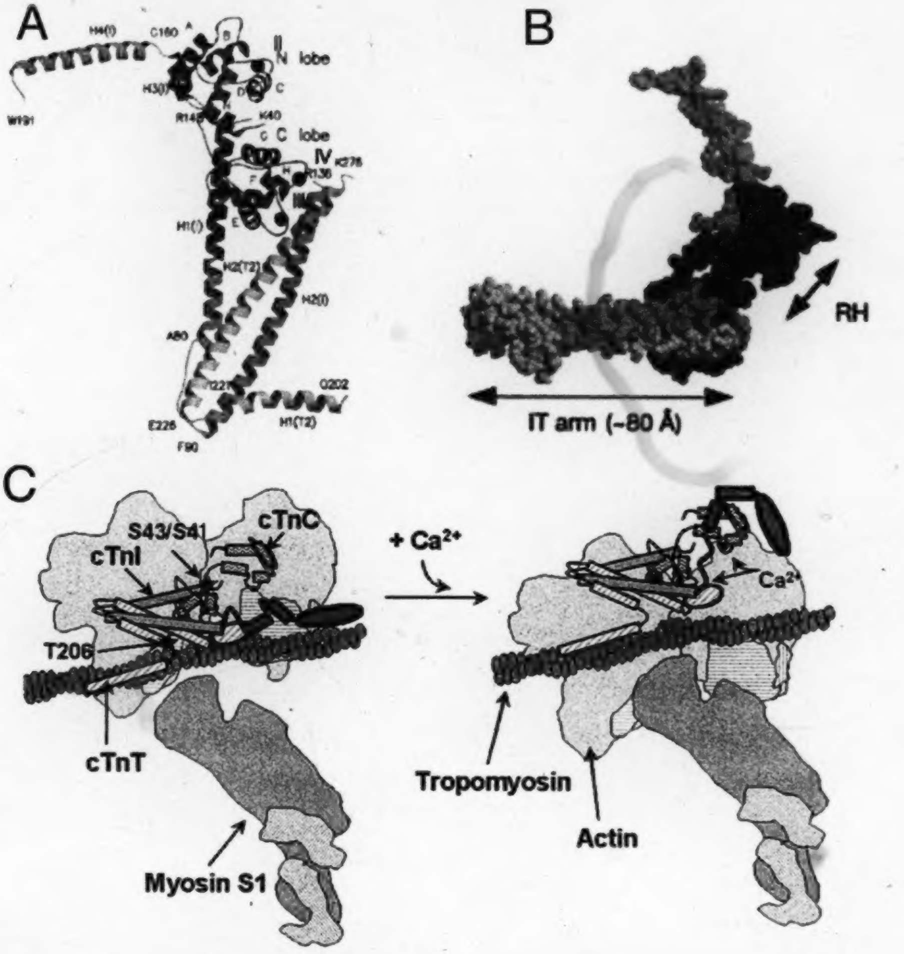

A
H4(I)
C160
A
B
II
N lobe
H3(I)
D
C
W191
R145
K40
G
C lobe
IV
K276
R136
H
F
H1(I)
E
H2(T2)
H2(I)
A80
T221
Q202
E226
H1(T2)
F90
B
RH
IT arm (~80 Å)
C
S43/S41
cTnC
cTnI
+ Ca2+
Ca2+
T206
cTnT
Tropomyosin
Actin
Myosin S1

Figure I-15. **Structure of the 'core' troponin complex**. (A) Ribbon diagram of the 52kDa 'core' complex where cTnC is shown in red, cTnI in blue, cTnT in yellow and Ca^{2+}-ions as black spheres (1J1E.pdb). (B) Space filling diagram of the same color scheme depicting the IT arm and the regulatory (RH) head. (C) Model of thin filament re-arrangement of the troponin complex during Ca^{2+}-induced cardiac contraction, freeing actin binding sites for myosin S1-ATPase. Figures and legend adapted from (*11, 51*).

THESIS GOALS AND PROJECT RATIONALE

The inhibitory region of troponin I is critical for inhibition of cardiac contraction. In spite of the given wealth of structural information currently available for cTnI interactions, a clear definition of its structural role and specifics of functionality is lacking. As previously mentioned, the inhibitory region is highly conserved throughout species and isoforms and does not tolerate mutations, as evidenced by the FHC mutation R144G. Likewise, modulation of cardiac contractility via PKC phosphorylation of Thr142 within this region makes it both an attractive candidate for further structural studies, as well as *in vitro* thermodynamic binding studies. The inhibitory region is an attractive target for directed structural studies by NMR spectroscopy due to its high solubility, and ease in production as a synthetic peptide. The majority of this thesis will deal in addressing seven main goals, which include structural aspects of this region in regards to competitive binding, mutation(s), modification(s), and will try to equate structural data to physiological function, in regards to cTnI as an individual entity and the troponin complex as a whole within the thin filament. This thesis work was started in the spring of 2000 in the laboratory of Dr. Brian D. Sykes in the Department of Biochemistry at the University of Alberta. As such this work was initiated prior to the published crystal structure of the 'core' troponin complex and various other biophysical studies. Therefore results will be compared and contrasted against all relevant data that both preceded and superseded publication of portions of this thesis.

Specific Aim #1 – Production of labeled inhibitory region for NMR structural studies

Chapter II - In attempting to solve the structure of the bound inhibitory region to cTnC, it was found that isotopically labeled cIp peptide (i.e. ^{13}C and/or ^{15}N) was a requirement for the unambiguous assignment on inter-domain NOE's, and as additional structural restraints for cIp. Unfortunately, synthetic production involving ^{15}N and/or ^{13}C was not cost effective and therefore an alternative approach to produce this labeled domain was required. Recombinant bacterial expression using exogenously added NMR labeled nuclei was the rational choice to pursue, however there was a severe lack of high-level expression peptide production systems available within the scientific community at the time of this project start. *Therefore for NMR structural studies of the inhibitory region of troponin I to begin; an efficient expression system designed for the production of isotopically labeled peptides needed to first be devised.*

Specific Aim #2 – Solve the NMR solution structure of the bound inhibitory region

Chapter III - Prior to my arrival in the laboratory, it was shown via {^{1}H, ^{15}N}-Heteronuclear Single Quantum Coherence (HSQC) NMR spectroscopy that the inhibitory region of cTnI binds preferentially to the C-lobe of cTnC (*73*). Since NMR structural studies are generally limited to the total molecular weight of a compound due to issues of line broadening and spectra crowding, it was decided to pursue the NMR solution structure of the inhibitory region bound solely to the C-lobe of cTnC in an attempt to reduce the total molecular weight of the complex. The elucidation of this binary structure could prove to be instrumental in understanding the structural aspects and functioning of this region in cardiac contraction. *The second goal of this thesis was therefore to determine the NMR solution structure of the inhibitory region bound to the C-lobe of cTnC.*

Specific Aim #3 – Electrostatic and thermodynamic consequences on complex formation

Chapter IV - Following the successful completion of the structure of the inhibitory region bound to the C-lobe, functional studies correlating mutation and modification with this domain could be addressed via NMR. Scanning glycine mutagenesis of the inhibitory region is implemented in attempting to mimic the known FHC mutation R144G. An exhaustive analysis of thermodynamic data with respect to electrostatics will show that the inhibitory region associates with cTnC via charge attraction. *The third goal of this thesis was to study the effects of mutations on the inhibitory region such as the FHC genotype R144G, and the effects of PKC phosphorylation of Thr142, on binding with the C-lobe of cTnC.*

Specific Aim #4 – The effects of cRP40 and the inhibitory region on binding to the C-lobe

Chapter V - Although the inhibitory region is known to bind to the C-lobe, there are other domains that additionally interact with the C-lobe, such as the cRP40 region of cTnI. It has been previously argued that the cRP40 region will displace the inhibitory region off the C-lobe when the two domains are in competition, indicating a common binding site (77). Additionally, cRP40 phosphorylation (Ser41 and Ser43) by PKC will be investigated to monitor the effects of phosphorylation as a regulator of cardiac contraction. *The fourth goal of this thesis was to probe via ^{15}N-NMR relaxation studies the nature of the interaction of the C-lobe of cTnC with both the inhibitory and cRP40 regions in an attempt to conclude the binding order and preferences, and correlate phosphorylation data of Ser41 and Ser43 with functional data.*

Specific Aim #5 – Interactions of cTnI with the N-lobe of cTnC

Chapter VI – An investigation of the cTnI interactions with the N-lobe is pursued to obtain a greater understanding of cTnI-cTnC interactions. The 'switch' region of cTnI is known to be the trigger of calcium-induced contraction via association with the N-lobe (see Figure I-12). The switch region is directly C-terminal to the inhibitory region on cTnI, as well contains a FHC mutation R161W, and contains a p21-activated kinase phosphorylation site at residue Ser149. Investigation of the interactions of the inhibitory region with the N-lobe will be investigated to determine if cIp binding is restricted solely to the C-lobe. *The fifth goal of this thesis was to study the effects of the inhibitory and switch regions on the N-lobe of cTnC and the effects of R161W and Ser149 phosphorylation on the mode of binding, in an attempt to gain insights into domain orientations during cardiac contraction.*

Specific Aim #6 – NMR studies of the full length troponin complex

Chapter VII - Over the years (as well within this thesis), structural studies of the troponin complex has focused on individual domains and have shied away from structural studies of the full length troponin complex due to size limitations and solubility concerns at concentrations required for NMR studies. With the availability of an ultra-high field 800 MHz NMR spectrometer at the National High Field Nuclear Magnetic Resonance Center (NANUC) situated at the University of Alberta, coupled with new NMR pulse sequence technology such as Transverse Relaxation Optimized Spectroscopy (TROSY), it has become possible to investigate high molecular weight complexes such as the cardiac troponin complex (38kDa). NMR studies of the whole complex can help to elucidate greater structural aspects of the troponin complex in its role in cardiac contraction, and provide dynamic insights as a supplement of the published crystal structure to determine cTnI flexibility. *The sixth goal of this thesis was to obtain NMR data of troponin I within a 38kDa troponin complex in an attempt to examine both dynamic and structural data of the troponin I subunit within the complex.*

Specific Aim #7 – Correlating structural studies with functional studies

Chapter VIII – Given the tremendous amount of functional, biophysical and structural information available regarding muscular contraction, concluding remarks are offered regarding the state of where we are in our understanding of muscular contraction and how far we still have to progress before a comprehensive understanding is available. This chapter will deal with addressing the results offered in the first seven chapters of this thesis, and how they have been instrumental in producing a greater understanding of cardiac contraction. As well, the recent crystal structure of a skeletal troponin complex and small angle neutron scattering data will be addressed, both of which differ significantly from the Takeda/Maeda crystal structure for the cardiac isoform (1J1D.pdb).

REFERENCES

1. Perry, S. V. (1999) *Mol Cell Biochem 190*, 9-32.
2. Perry, S. V. (2003) *J Muscle Res Cell Motil 24*, 593-6.
3. Tobacman, L. S. (1996) *Annu. Rev. Physiol. 58*, 447-481.
4. Farah, C. S., and Reinach, F. C. (1995) *FASEB J. 9*, 755-767.
5. Squire, J. M., and Morris, E. P. (1998) *FASEB Journal 12*, 761-771.
6. Filatov, V. L., Katrukha, A. G., Bulargina, T. V., and Gusev, N. B. (1999) *Biochemistry (Mosc). 64*, 969-85.
7. Clark, K. A., McElhinny, A. S., Beckerle, M. C., and Gregorio, C. C. (2002) *Annu Rev Cell Dev Biol 18*, 637-706.
8. Gordon, A. M., Homsher, E., and Regnier, M. (2000) *Physiological Reviews 80*, 853-924.
9. Solaro, R. J. (2003) *Adv Exp Med Biol 538*, 389-401; discussion 401-2.
10. Solaro, R. J. (1999) *Circ Res 84*, 122-4.
11. Kobayashi, T., and Solaro, R. J. (2004) *Annu Rev Physiol.*
12. Sumandea, M. P., Burkart, E. M., Kobayashi, T., De Tombe, P. P., and Solaro, R. J. (2004) *Ann N Y Acad Sci 1015*, 39-52.
13. Perry, S. V. (1999) *Mol. Cell. Biochem. 190*, 9-32.
14. Ji, L. L., Fu, R. G., Mitchell, E. W., Griffiths, M., Waldrop, T. G., and Swartz, H. M. (1994) *Acta Physiol Scand 151*, 279-90.
15. Biofluids-Laboratory. (2005) *Cardiovascular Physiology*, http://www.eng.uab.edu/bme/faculty/aanayiot/biofluids/cardiova.htm.
16. Johnson, G. J. (2005) *The World of Biology*, http://www.sirinet.net/~jgjohnso/heartstructure.html.
17. Voet, D., Voet, J. G., and Pratt, C. W. (1999) *Fundamentals of Biochemistry*, 180-186, John Wiley & Sons, Inc., Toronto.
18. Klabunde, R. E. (2004) *Cardiovascular Physiology Concepts*, Chapter 4, Linppincott Williams & Wilkins.
19. Hanson, J., and Huxley, H. E. (1953) *Nature 172*, 530-2.
20. Huxley, H. E. (1953) *Biochim Biophys Acta 12*, 387-94.
21. Huxley, H. E. (1953) *Proc R Soc Lond B Biol Sci 141*, 59-62.
22. Martini. (2003) *Fundamentals of Anatomy & Physiology, 6th Edition, Chapter 10*, Prentice Hall.
23. Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F., and Holmes, K. C. (1990) *Nature 347*, 37-44.
24. Holmes, K. C., Popp, D., Gebhard, W., and Kabsch, W. (1990) *Nature 347*, 44-9.
25. Herzberg, O., and James, M. N. G. (1988) *J. Mol. Biol. 203*, 761-779.
26. Herzberg, O., and James, M. N. G. (1985) *Nature 313*, 653-659.
27. Kasai, M., Yamaguchi, N., and Kawasaki, T. (1995) *J Biochem (Tokyo) 117*, 251-6.
28. Kawasaki, T., and Kasai, M. (1994) *Biochem Biophys Res Commun 199*, 1120-7.
29. Meissner, G., Rousseau, E., Lai, F. A., Liu, Q. Y., and Anderson, K. A. (1988) *Mol Cell Biochem 82*, 59-65.
30. Anderson, K., and Meissner, G. (1995) *J Gen Physiol 105*, 363-83.
31. Head, J. F., and Perry, S. V. (1974) *Biochemical Journal 137*, 145-154.

32. Sia, S. K., Li, M. X., Spyracopoulos, L., Gagne, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) *J Biol Chem 272*, 18216-21.
33. Xu, C., Craig, R., Tobacman, L., Horowitz, R., and Lehman, W. (1999) *Biophys. J. 77*, 985-992.
34. Fivth, E. M. M. *fiVth element Medical Media*, http://www.fiVth.com.
35. Kawasaki, H., and Kretsinger, R. H. (1994) *Protein Profile 1*, 343-517.
36. Kawasaki, H., and Kretsinger, R. H. (1995) *Protein Profile 2*, 297-490.
37. Gagne, S. M., Li, M. X., McKay, R. T., and Sykes, B. D. (1998) *Biochem Cell Biol 76*, 302-12.
38. Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry 36*, 4386-4392.
39. Ikura, M. (1996) *Trends Biochem Sci 21*, 14-7.
40. Babu, A., Su, H., Ryu, Y., and Gulati, J. (1992) *J Biol Chem 267*, 15469-74.
41. Brito, R. M., Krudy, G. A., Negele, J. C., Putkey, J. A., and Rosevear, P. R. (1993) *J Biol Chem 268*, 20966-73.
42. Trigo-Gonzalez, G., Awang, G., Racher, K., Neden, K., and Borgford, T. (1993) *Biochemistry 32*, 9826-31.
43. Liu, W., Dotson, D. G., Lin, X., III, J. J. M., Gonzalez-Garay, M. L., Lu, Q., and Putkey, J. A. (1994) *FEBS Lett. 347*, 152-156.
44. Hazard, A. L., Kohout, S. C., Stricker, N. L., Putkey, J. A., and Falke, J. J. (1998) *Protein Sci 7*, 2451-9.
45. Dotson, D. G., and Putkey, J. A. (1993) *J Biol Chem 268*, 24067-73.
46. Sweeney, H. L., Brito, R. M., Rosevear, P. R., and Putkey, J. A. (1990) *Proc Natl Acad Sci U S A 87*, 9538-42.
47. Finley, N., Dvoretsky, A., and Rosevear, P. R. (2000) *J Mol Cell Cardiol 32*, 1439-46.
48. Finley, N. L., Howarth, J. W., and Rosevear, P. R. (2004) *Biochemistry 43*, 11371-9.
49. Robertson, S. P., Johnson, J. D., and Potter, J. D. (1981) *Biophys J 34*, 559-69.
50. Sia, S. K., Li, M. X., Spyracopoulos, L., Gagné, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) *J. Biol. Chem. 272*, 18216-18221.
51. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
52. Malnic, B., Farah, C. S., and Reinach, F. C. (1998) *J Biol Chem 273*, 10594-601.
53. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.
54. Spyracopoulos, L., Li, M. X., Sia, S. K., Gagné, S. M., Chandra, M., Solaro, R. J., and Sykes, B. D. (1997) *Biochemistry 36*, 12138-12146.
55. Luo, Y., Wu, J.-L., Gergely, J., and Tao, T. (1997) *Biochemistry 36*, 11027-11035.
56. Tao, T., Gong, B. J., Grabarek, Z., and Gergely, J. (1999) *Biochim Biophys Acta 1450*, 423-33.
57. Maytum, R., Geeves, M. A., and Lehrer, S. S. (2002) *J Biol Chem 277*, 29774-80.
58. Heeley, D. H., Golosinska, K., and Smillie, L. B. (1987) *J. Biol. Chem. 262*, 9971-9978.
59. Pearlstone, J. R., and Smillie, L. B. (1985) *Can J Biochem Cell Biol 63*, 212-8.
60. Redwood, C. S., Moolman-Smook, J. C., and Watkins, H. (1999) *Cardiovasc Res 44*, 20-36.

61. Gomes, A. V., Venkatraman, G., Davis, J. P., Tikunova, S. B., Engel, P., Solaro, R. J., and Potter, J. D. (2004) *J Biol Chem 279*, 49579-87.
62. Potter, J. D., Sheng, Z., Pan, B. S., and Zhao, J. (1995) *J Biol Chem 270*, 2557-62.
63. MacGeoch, C., Barton, P. J., Vallins, W. J., Bhavsar, P., and Spurr, N. K. (1991) *Hum Genet 88*, 101-4.
64. Bhavsar, P. K., Brand, N. J., Yacoub, M. H., and Barton, P. J. (1996) *Genomics 35*, 11-23.
65. Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) *J Mol Biol 271*, 728-50.
66. Sykes, B. D. (2003) *Nat Struct Biol 10*, 588-9.
67. Solaro, R. J., el-Saleh, S. C., and Kentish, J. C. (1989) *Mol Cell Biochem 89*, 163-7.
68. Talbot, J. A., and Hodges, R. S. (1979) *J. Biol. Chem. 254*, 3720-3723.
69. Talbot, J. A., and Hodges, R. S. (1981) *J. Biol. Chem. 256*, 2798-2802.
70. Syska, H., Wilkinson, J. M., Grand, R. J., and Perry, S. V. (1976) *Biochem J 153*, 375-87.
71. Sasse, S., Brand, N. J., Kyprianou, P., Dhoot, G. K., Wade, R., Arai, M., Periasamy, M., Yacoub, M. H., and Barton, P. J. (1993) *Circ Res 72*, 932-8.
72. Vassylyev, D. G., Takeda, S., Wakatsuki, S., Maeda, K., and Maeda, Y. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 4847-4852.
73. Li, M. X., Spyracopoulos, L., Beier, N., Putkey, J. A., and Sykes, B. D. (2000) *Biochemistry 39*, 8782-8790.
74. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1998) *Biophys. J. 74*, A51.
75. Li, M. X., Wang, X., Lindhout, D. A., Buscemi, N., Van Eyk, J. E., and Sykes, B. D. (2003) *Biochemistry 42*, 14460-8.
76. Mercier, P., Ferguson, R. E., Irving, M., Corrie, J. E., Trentham, D. R., and Sykes, B. D. (2003) *Biochemistry 42*, 4333-48.
77. Mercier, P., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 2902-11.
78. Mercier, P., Spyracopoulos, L., and Sykes, B. D. (2001) *Biochemistry 40*, 10063-77.
79. McKay, R. T., Pearlstone, J. R., Corson, D. C., Gagne, S. M., Smillie, L. B., and Sykes, B. D. (1998) *Biochemistry 37*, 12419-30.
80. McKay, R. T., Saltibus, L. F., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 12731-12738.
81. McKay, R. T., Tripet, B. P., Hodges, R. S., and Sykes, B. D. (1997) *J Biol Chem 272*, 28494-500.
82. McKay, R. T., Tripet, B. P., Pearlstone, J. R., Smillie, L. B., and Sykes, B. D. (1999) *Biochemistry 38*, 5478-89.
83. Wang, X., Li, M. X., Spyracopoulos, L., Beier, N., Chandra, M., Solaro, R. J., and Sykes, B. D. (2001) *J Biol Chem 276*, 25456-66.
84. Wang, X., Li, M. X., and Sykes, B. D. (2002) *J Biol Chem 277*, 31124-33.
85. Tripet, B., De Crescenzo, G., Grothe, S., O'Connor-McCourt, M., and Hodges, R. (2002) *J Mol Biol 323*, 345.
86. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.
87. Lindhout, D. A., and Sykes, B. D. (2003) *J Biol Chem 278*, 27024-34.

88. Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) *FEBS Lett 453*, 107-12.
89. D'Costa, M., Fleming, E., and Patterson, M. C. (1997) *Am J Clin Pathol 108*, 550-5.
90. Dargis, R., Pearlstone, J. R., Barrette-Ng, I., Edwards, H., and Smillie, L. B. (2002) *J Biol Chem 277*, 34662-5.
91. Kobayashi, T., Leavis, P. C., and Collins, J. H. (1996) *Biochimica et Biophysica Acta 1294*, 25-30.
92. Ngai, S.-M., and Hodges, R. S. (1992) *J. Biol. Chem. 267*, 15715-15720.
93. Cachia, P. J., Sykes, B. D., and Hodges, R. S. (1983) *Biochemistry 22*, 4145-52.
94. Campbell, A. P., and Sykes, B. D. (1991) *J Mol Biol 222*, 405-21.
95. Campbell, A. P., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) *Biochim Biophys Acta 1160*, 35-54.
96. Ngai, S.-M., Sönnichsen, F. D., and Hodges, R. S. (1994) *J. Biol. Chem. 269*, 2165-2172.
97. Hernandez, G., Blumenthal, D. K., Kennedy, M. A., Unkefer, C. J., and Trewhella, J. (1999) *Biochemistry 38*, 6911-7.
98. Tung, C. S., Wall, M. E., Gallagher, S. C., and Trewhella, J. (2000) *Protein Sci 9*, 1312-26.
99. Brown, L. J., Sale, K. L., Hills, R., Rouviere, C., Song, L., Zhang, X., and Fajer, P. G. (2002) *Proc Natl Acad Sci U S A 99*, 12765-70.
100. Hernandez, O. M., Housmans, P. R., and Potter, J. D. (2001) *J Appl Physiol 90*, 1125-36.
101. McDonough, J. L., Arrell, D. K., and Van Eyk, J. E. (1999) *Circ Res 84*, 9-20.
102. Murphy, A. M., Kogler, H., Georgakopoulos, D., McDonough, J. L., Kass, D. A., Van Eyk, J. E., and Marban, E. (2000) *Science 287*, 488-91.
103. Van Eyk, J. E., Powers, F., Law, W., Larue, C., Hodges, R. S., and Solaro, R. J. (1998) *Circ Res 82*, 261-71.
104. McDonough, J. L., Labugger, R., Pickett, W., Tse, M. Y., MacKenzie, S., Pang, S. C., Atar, D., Ropchan, G., and Van Eyk, J. E. (2001) *Circulation 103*, 58-64.
105. Lee, T. H., and Goldman, L. (2000) *N Engl J Med 342*, 1187-95.
106. Smith, J. J. (1980) *Circulatory Physiology: The Essentials*, Baltimore, Md: Waverly Press.
107. Solaro, R. J., Moir, A. J., and Perry, S. V. (1976) *Nature 262*, 615-7.
108. Katz, A. M. (1990) *Eur Heart J 11 Suppl A*, 27-31.
109. Metzger, J. M., and Westfall, M. V. (2004) *Circ Res 94*, 146-58.
110. Noland, T. A., Jr., Guo, X., Raynor, R. L., Jideama, N. M., Averyhart-Fullard, V., Solaro, R. J., and Kuo, J. F. (1995) *J Biol Chem 270*, 25445-54.
111. Noland, T. A., Jr., Raynor, R. L., Jideama, N. M., Guo, X., Kazanietz, M. G., Blumberg, P. M., Solaro, R. J., and Kuo, J. F. (1996) *Biochemistry 35*, 14923-31.

CHAPTER II

Design of a peptide expression system for NMR studies of the troponin system

OVERVIEW

This chapter addresses a void in the molecular biology community with regards to a high-yield expression system for peptide production, with an emphasis on the ability to fully incorporate NMR active nuclei for structural studies. During my thesis research of the troponin complex and the inhibitory region of cTnI, I required milligram quantities of labeled (i.e. ^{13}C and/or ^{15}N) inhibitory region for NMR structural elucidation. Unfortunately the only available protein source was synthetic, unlabeled peptide from the Alberta Peptide Institute. Production of a labeled synthetic peptide would cost in excess of $20,000 and was not a viable route to pursue. The only available option was to initiate a protocol for peptide production via recombinant expression. I entered into a collaborative effort with the laboratory of Dr. G. Marius Clore from the National Institutes of Health in Bethesda, Maryland, and have been extremely successful in the production of labeled peptides for use in NMR studies via a fusion protein technique with the B1 domain of protein G from *Streptococcus*. This work was published in the August 2003 issue of *Protein Science* [Protein Science. 2003 Aug; **12(8)**:1786-91] and has been extremely well received by the academic community, with our laboratory routinely being asked weekly for the expression vector and details of the protocol (*1*).

Fusion protein constructs of the 56 amino acid globular protein GB1 with various peptide sequences, coupled with the incorporation of a histidine tag for affinity purification, have generated high yield fusion protein constructs. Methionine residues were inserted into the constructs to generate pure peptides following CNBr cleavage, yielding a system that is efficient and cost effective for isotopic labeling of peptides for NMR studies, and other disciplines such as mass spectrometry. Six peptides of varying sequences and hydrophobicities were expressed using this GB1 fusion protein technique, and produced soluble fusion protein constructs in all cases. The ability to easily express and purify recombinant peptides in high yields is applicable for biomedical research, medicinal and pharmaceutical applications.

INTRODUCTION

Polypeptides represent a significant group of bio-molecules, which can play an important role in various cellular functions. Included in this group of bio-molecules is the 51 residue polypeptide insulin, which initiates a hormonal cellular signaling cascade in response to blood-glucose concentrations (*2*). Implications for polypeptides to behave as independent domains are displayed by the 26 residue α-helical lytic bee venom polypeptide melittin (*3*), as well as an effector of intracellular Ca^{2+}-signaling; phospholamban, an integral membrane 52 residue polypeptide composed of two α-helical regions connected by a β-turn (*4, 5*). Polypeptides attributed to diseased states includes the 37 residue islet amyloid polypeptide (hIAPP), which has been found to accumulate as amyloid fibrils in the pancreas of individuals with type II diabetes (*6, 7*), as well polypeptides such as LL-37 has demonstrated strong anti-microbial properties, leading to a novel class of treatment remedies of antibiotic resistant pathogens (*8, 9*). Polypeptide regions can be generated from larger multidomain proteins using various proteolytic procedures, with generated regions retaining full biological activity of the parent domains. Bio-medical advancements have exploited this ability of peptides to retain activity to allow peptides to be used as tools in research, such as for potential therapeutic agents against disease (*10, 11*) and use as peptide vaccines (*12*).

The current peptide production technique commonly employed is solid-state synthesis, yet this can be hindered by high cost and relatively low yields, as well as the inability to accurately produce peptides of lengths of over approximately 50 residues with acceptable yields. Furthermore, the ability to introduce isotopic labels (^{13}C, ^{15}N, ^{2}H) within polypeptides by solid-state methodology for NMR studies is not a viable option, as the costs involved quickly escalate. The ability to recombinantly produce peptides is an alternative to solid-state synthesis, as cell culture expression can yield high production at relatively low cost, with easy incorporation of isotopic labels. Unfortunately, expression of peptides *in vivo* has met with limited success as peptides are relatively poorly expressed in cell culture; attributed to an unstructured state in solution susceptible to cellular proteases, as well as solubility concerns of the expressed peptide. Peptide production may also produce highly toxic effects on cells during expression (*13*).

One solution to high yield recombinant peptide production is the use of a fusion protein construct, which helps in peptide stability and solubility following expression. Fusion proteins offer protection of expressed peptides from cellular proteases, incorporation of various affinity tags for ease in fusion protein purification and introduction of proteolytic/chemical cleavage sites for generation of subsequent peptide fragments. Several groups have reported success with various fusion protein constructs, with many using a hydrophobic fusion protein construct which produces a construct which is sequestered within inclusion bodies (*13-15*). However, reconstitution of inclusion body proteins provides additional work and use of chemical denaturants for subsequent peptide production.

In this chapter, I will report the production of six different fusion protein constructs using the B1 immunoglobulin binding domain of *Streptococcal* protein G (*16*) termed GEV1 vector (*17-19*), a 56 amino acid soluble globular domain combined with a poly-histidine tag affinity purification protocol. The advantage of the GB1 fusion approach over other previous techniques is the production of a soluble fusion protein construct, with high expression yields and ease of purification. The ability of GB1 to express multiple differing peptide fragments of various composition and length is demonstrated in this study. I have cloned, expressed and purified six different peptides. These include cIp-RR-20; a 20 residue peptide of human cardiac troponin I ($\text{cTnI}_{128\text{-}147}$) which plays a role in regulation of heart muscle contractility (*20-22*), as well as two mutants of cIp-RR-20 (cIp-RG-20 and cIp-GR-20) which yield altered activities, with cIp-GR-20 being associated with the heart disease familial hypertrophic cardiomyopathy (*23*, *24*). I have also produced and purified TRTK-12; a 12 residue homology peptide of the actin capping protein ($\text{CapZ}\alpha 1_{265\text{-}276}$) which has been shown to interact with the calcium binding protein S-100 (*25*), FLQS-26; a 26 residue peptide from human sodium proton exchanger isoform 1 ($NHE1_{155-179}$) which constitutes a predicted hydrophobic transmembrane helix (*26, 27*), and EDQL-26; a 26 residue peptide of cardiac troponin T ($cTnT_{226-251}$) which makes a helical coiled-coil interaction with cardiac troponin I (*28*). Using synthetic DNA oligonucleotides to code for the peptides of interest, as well as the introduction of methionine residues for subsequent cyanogen bromide (CNBr) cleavage following fusion protein purification, I have demonstrated a system which can be implemented, cloned, expressed and purified in as little as two weeks with high yields and minimal effort. Moreover, I demonstrate the ability of this system to produce isotopic labeling of peptides for NMR studies and provide a strong prospect for future production of peptides for various industrial, pharmaceutical and general research applications.

EXPERIMENTAL PROCEDURES

II-A. GEV-1 vector and primer design

Design of the GEV1 plasmid is shown in Figure II-1. Sense and antisense DNA oligonucleotide primers were synthesized using optimized codons for *E. coli* K12 based on the primary amino acid sequence for each of the 6 peptides for insertion into the GEV1 multiple cloning-site. Primer design for all peptides is detailed in Figure II-2. The codon biases of *E. coli* K12 are shown in Figure II-3, with a more detailed description of usage in Appendix A-1. In addition to coding for amino acid sequences, a methionine codon was introduced on either side of the peptide sequence, a SmaI cut site was introduced on the 3' end of the sense strand and NheI /XhoI cut sites were introduced for cloning into GEV1 vector. Sense and antisense oligonucleotides were annealed together at 100°C for 1 hour (100mM Tris HCl (pH 7.6), 1M NaCl, 10mM EDTA), followed by ethanol precipitation (2.5 volumes of 100% Ethanol, 1 hour at -20°C) and re-suspension in distilled, deionized water (ddH_2O). GEV1 vector was simultaneously digested with NheI/XhoI, agarose gel purified and ligated to individual annealed oligonucleotides. Ligation was performed at an insert to vector ratio of 3:1 and incubated overnight at 16°C with T4 DNA ligase (250mM Tris HCl (pH 7.6), 50mM $MgCl_2$, 5mM ATP and 5mM DTT). Ligated solutions were transformed directly into *E. coli* cloning strain XL1 Blue via heat shock protocol (42°C for 90 seconds followed by ice incubation for 2 minutes), and plated onto 2xYT/Amp plates and incubated 17 hours at 37°C. Colonies were individually selected, added to 5mL cultures of 2xYT/Amp and incubated 17 hours at 37°C with shaking. Following plasmid isolation (Bio 101 DNA Mini-Prep kit), positive colonies were screened against SmaI restriction digests and respective vectors were transformed (heat shock) into *E. coli* expression strain BL21 DE3 (pLysS).

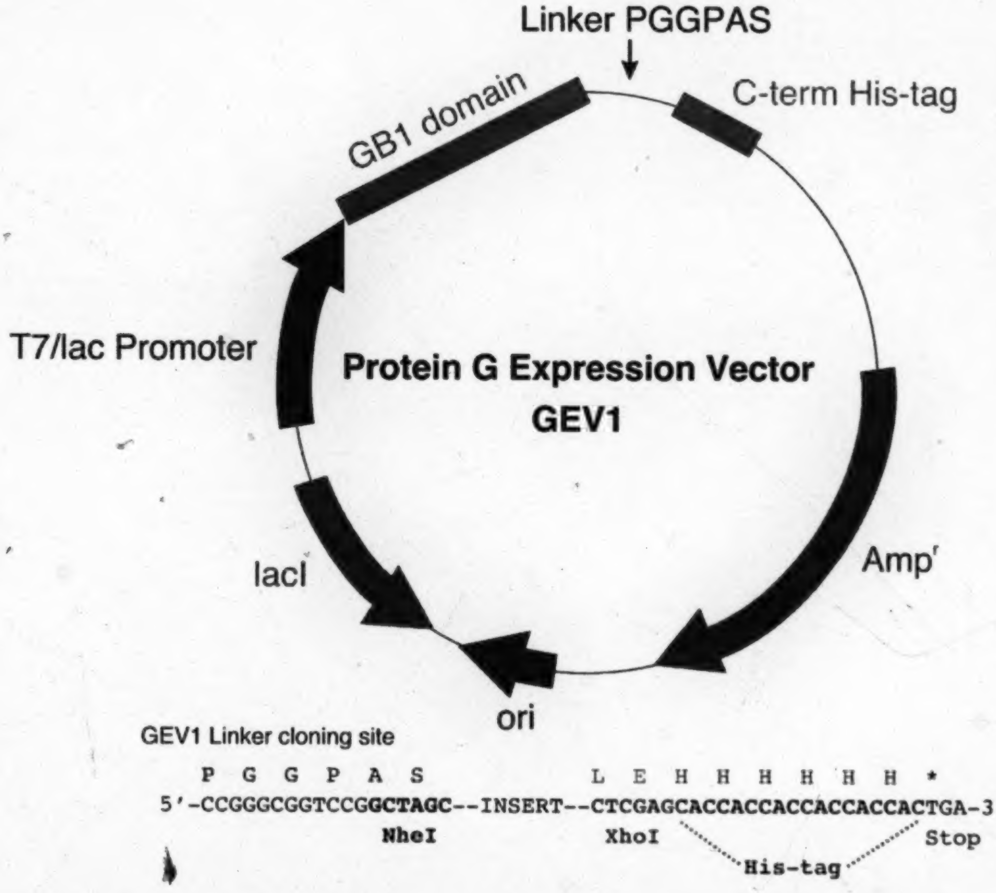
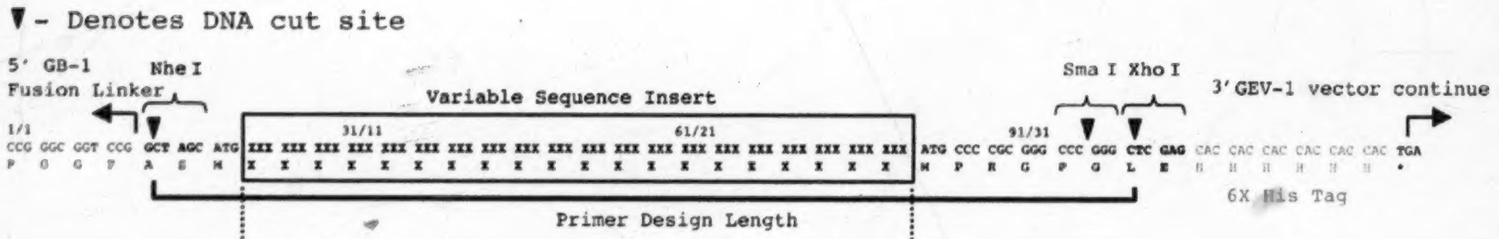

Linker PGGPAS
GB1 domain
C-term His-tag
T7/lac Promoter
Protein G Expression Vector
GEV1
lacI
Amp^r
ori
GEV1 Linker cloning site
P G G P A S L E H H H H H H *
5'-CCGGGCGGTCCGGCTAGC--INSERT--CTCGAGCACCACCACCACCACCACTGA-3'
NheI
XhoI
His-tag
Stop

Figure II-1. **GEV1 vector map**. A schematic diagram of the GEV1 plasmid (pET21a), which contains ampicillin resistance. The GB1 domain is under the influence of a T7 promoter. The linker region is shown below the vector, with the multiple cloning-site containing NheI and XhoI cut sites for ligation. Figure adapted from (*18*).


▼ - Denotes DNA cut site
5′ GB-1
Fusion Linker
Nhe I
Variable Sequence Insert
Sma I Xho I
3′ GEV-1 vector continue
1/1
31/11
61/21
91/31
CCG GGC GGT CCG GCT AGC ATG XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX XXX ATG CCC CGC GGG CCC GGG CTC GAG CAC CAC CAC CAC CAC CAC TGA
P G G P A S M X X X X X X X X X X X X X X X X X X X X M P R G P G L E H H H H H H •
Primer Design Length
6X His Tag

Primer name	Sense Primer Sequence
cIp-RR-20	ACC CAG AAA ATT TTT GAT CTG CGC GGC AAA TTT AAA CGC CCG ACC CTG CGC CGC GTG CGC T Q K I F D L R G K F K R P T L R R V R
cIp-GR-20	ACC CAG AAA ATT TTT GAT CTG CGC GGC AAA TTT AAA CGC CCG ACC CTG GGC CGC GTG CGC T Q K I F D L R G K F K R P T L G R V R
cIp-RG-20	ACC CAG AAA ATT TTT GAT CTG CGC GGC AAA TTT AAA CGC CCG ACC CTG CGC GGC GTG CGC T Q K I F D L R G K F K R P T L R G V R
TRTK-12	ACC CGC ACC AAA ATT GAT TGG AAC AAA ATT CTG AGC T R T K I D W N K I L S
EDQL-26	GAG GAT CAG CTG CGC GAG AAA GCT AAA GAG CTG TGG CAG AGC ATT TAT AAT CTG GAG GCT GAG AAA TTT GAT CTG CAG E D Q L R E K A K E L W Q S I Y N L E A E K F D L Q
FLQS-26	TTT CTG CAG AGC GAT GTG TTT TTT CTG TTT CTG CTG CCG CCG ATT ATT CTG GAT GCG GGC TAT TTT CTG CCG CTG CGC F L Q S D V F F L F L L P P I I L D A G Y F L P L R

Figure II-2. **Design of peptide DNA oligonucleotides for insertion into the GEV1 vector**. All primers listed as only sense strands (5' to 3'), with coded amino acid sequence shown below. All primers contain a 5'-NheI cut site and 3'-SmaI/XhoI cut sites and a N- and C-terminal Met residue for CNBr cleavage of the subsequent fusion protein following affinity purification.

```
Escherichia coli K12 [gbbct]: 5089 CDS's (1608122 codons)

fields: [triplet] [frequency: per thousand] ([number])

UUU F 0.57 22.4 ( 35982)  UCU S 0.15  8.5 ( 13687)  UAU Y 0.57 16.3 ( 26266)  UGU C 0.45  5.2 (  8340)
UUC F 0.43 16.6 ( 26678)  UCC S 0.15  8.6 ( 13849)  UAC Y 0.43 12.3 ( 19728)  UGC C 0.55  6.4 ( 10347)
UUA L 0.13 13.9 ( 22376)  UCA S 0.12  7.2 ( 11511)  UAA * 0.64  2.0 (  3246)  UGA * 0.29  0.9 (  1468)
UUG L 0.13 13.7 ( 22070)  UCG S 0.15  8.9 ( 14379)  UAG * 0.07  0.2 (   378)  UGG W 1.00 15.3 ( 24615)

CUU L 0.10 11.0 ( 17754)  CCU P 0.16  7.1 ( 11340)  CAU H 0.57 12.9 ( 20720)  CGU R 0.38 21.0 ( 33694)
CUC L 0.10 11.0 ( 17723)  CCC P 0.13  5.5 (  8915)  CAC H 0.43  9.7 ( 15595)  CGC R 0.40 22.0 ( 35306)
CUA L 0.04  3.9 (  6212)  CCA P 0.19  8.5 ( 13707)  CAA Q 0.35 15.4 ( 24835)  CGA R 0.06  3.6 (  5716)
CUG L 0.50 52.7 ( 84673)  CCG P 0.52 23.2 ( 37328)  CAG Q 0.65 28.8 ( 46319)  CGG R 0.10  5.4 (  8684)

AUU I 0.51 30.4 ( 48818)  ACU T 0.17  9.0 ( 14397)  AAU N 0.45 17.7 ( 28465)  AGU S 0.15  8.8 ( 14092)
AUC I 0.42 25.0 ( 40176)  ACC T 0.43 23.4 ( 37624)  AAC N 0.55 21.7 ( 34912)  AGC S 0.28 16.1 ( 25843)
AUA I 0.07  4.3 (  6962)  ACA T 0.13  7.1 ( 11366)  AAA K 0.77 33.6 ( 54097)  AGA R 0.04  2.1 (  3337)
AUG M 1.00 27.7 ( 44614)  ACG T 0.27 14.4 ( 23124)  AAG K 0.23 10.2 ( 16401)  AGG R 0.02  1.2 (  1987)

GUU V 0.26 18.4 ( 29569)  GCU A 0.16 15.4 ( 24719)  GAU D 0.63 32.2 ( 51852)  GGU G 0.34 24.9 ( 40019)
GUC V 0.22 15.2 ( 24477)  GCC A 0.27 25.5 ( 40993)  GAC D 0.37 19.0 ( 30627)  GGC G 0.40 29.4 ( 47309)
GUA V 0.15 10.9 ( 17508)  GCA A 0.21 20.3 ( 32666)  GAA E 0.69 39.5 ( 63517)  GGA G 0.11  7.9 ( 12776)
GUG V 0.37 26.2 ( 42212)  GCG A 0.35 33.6 ( 53988)  GAG E 0.31 17.7 ( 28522)  GGG G 0.15 11.0 ( 17704)


Coding GC 51.80% 1st letter GC 58.89% 2nd letter GC 40.72% 3rd letter GC 55.79%
Note: first figure is proportional use for that amino acid. Genetic code 11: Eubacterial
```

Figure II-3. ***Escherichia coli*** **K12 codon bias**. Genetic code translation table indicating preferences of specific codons for K12. Codons with the highest probability of occurrence (bold) were used in primer design to ensure maximal expression of fusion proteins. Figure statistics taken from the NCBI (*29*).

II-B. Sequencing of GEV1 fusion plasmids

All ligated vectors confirmed positive by SmaI digest were subjected to further confirmation by DNA sequencing using the T7-terminator primer. The sequence of the T7-terminator primer used is 5'-TATGCTAGTTATTGCTCAG-3'. Sequencing results are listed in Appendix A-2.

II-C. Test Expression of GB1 fusion proteins

E. coli BL21 DE3 (pLysS) cells containing each respective GB1 fusion vector were grown in 50mL 2xYT containing 100*u*g/mL ampicillin overnight 37°C with shaking. 1 Liter cultures of 2xYT containing 100*u*g/mL ampicillin were inoculated with 5mL of overnight culture of each fusion protein construct, grown to OD_{600} 0.8 at 37°C with shaking, and induced with 1mM isopropyl β-D-thiogalactoside (IPTG) for 4 hours. To determine the extent of soluble expression, each fusion protein construct was pelleted, broken in a French pressure cell, cleared by centrifugation and supernatant analyzed on an 8-22% gradient SDS-PAGE gel (Mini-Protean III - Biorad).

II-D. Labeling of GB1 fusion proteins clp-RR-20 and TRTK-12

To prove the ability of incorporation of NMR active nuclei in recombinant expression of fusion proteins, clp-RR-20 and TRTK-12 were tested. Starter cultures were grown in 50mL 2xYT containing 100ug/mL ampicillin at 37°C until OD_{600} 0.8 was reached. For ^{15}N labeling of fusion proteins 1 Liter cultures of M9 minimal media (50mM Na_2HPO_4, 25mM KH_2PO_4, 10mM NaCl, pH 7.3) containing 50mL of ^{15}N-enriched media (Cambridge Isotope Laboratories, Inc.) and 1.1g ^{15}N-$(NH_4)_2SO_4$ as the sole nitrogen source and 10g of glucose, were inoculated with 5mL of starter culture, grown to OD_{600} 0.8 at 37°C, inoculated with 1mM IPTG, and allowed to induce for 4 hours prior to cell harvesting. For $^{13}C/^{15}N$ labeling of fusion proteins 1 Liter cultures of M9 minimal media containing 50mL of $^{13}C/^{15}N$-enriched media (Cambridge Isotope Laboratories, Inc.), 1.1g ^{15}N-$(NH_4)_2SO_4$ and 2.5g of ^{13}C-glucose.

II-E. Purification of GB1 fusion proteins

BL21 DE3 (pLysS) cells which expressed respective fusion proteins were pelleted, resuspended in 50mM Tris buffer (pH 8.0), broken in a French pressure cell, centrifuged and passed through a 0.22*u*m filter to remove all insoluble cellular debris. Supernatant was loaded onto a 25mL column of fast flow chelating sepharose, previously charged with 50mM $NiSO_4$, and equilibrated in 5mM imidazole, 500mM NaCl, 20mM TrisHCl, pH 7.9. The column was then washed with 40mM imidazole, 500mM NaCl, 10mM TrisHCl, pH 7.9, and was followed by fusion protein elution with 100mM EDTA, 500mM NaCl, 20mM TrisHCl, pH 7.9. Fusion protein eluent was lyophilized to dryness, desalted on a G25 sephadex column (10mM NH_4HCO_3, pH 8.0) and lyophilized to dryness.

II-F. CNBr cleavage of GB1 fusion proteins and purification of peptides

GB1 fusion proteins were dissolved in 0.1N HCl (5mg/mL), to which solid CNBr was added to a final molar concentration of 100:1, and left in the dark at room temperature for 24 hours. Solutions were then diluted 10 fold with distilled deionized H_2O (ddH_2O) and twice lyophilized to dryness. Samples were subjected to a reverse-phase HPLC column, individual peaks were collected with molecular weights verified by MALDI-TOF mass spectrometry. Peaks containing pure peptide were pooled and twice lyophilized to dryness to remove all organic solvents.

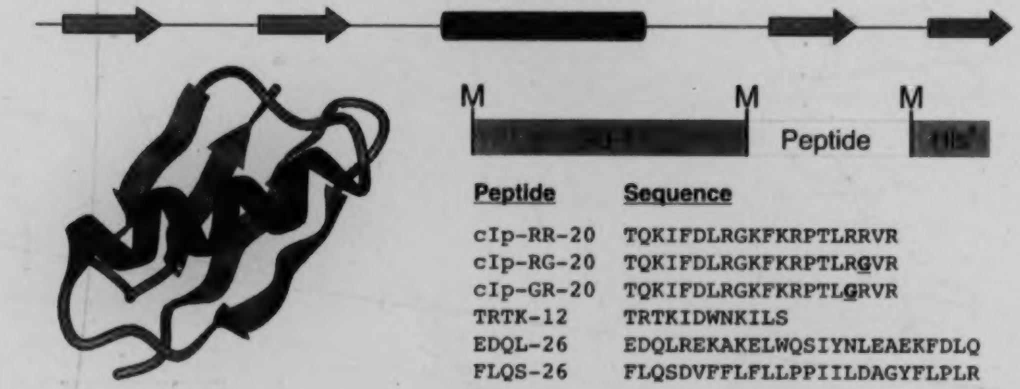
II-G. NMR spectroscopy

NMR experiments were performed on a Varian INOVA 500MHz NMR spectrometer equipped with a triple resonance probe and Z-axis pulsed field gradients. All recorded spectra were referenced to an external DSS standard. 5.3mg of ^{1}H-GB1_clp-RR-20_His was dissolved in 525*u*L of 0.1N HCl with 25*u*L of D_2O, to which a 100-fold excess of CNBr was added directly to the NMR tube. 1D-^{1}H Scans were taken at 15-minute intervals at 25°C for 24 hours. All one-dimensional spectra were processed and analyzed using vnmr (Varian associates). 2.5mg of ^{15}N-clp-RR-20 and 3.1 mg of ^{15}N/^{13}C-clp-RR-20 were dissolved separately into 500*u*L NMR buffer (100mM KCl, 10mM imidazole, 0.015% NaN_3, pH 6.7) with all two-dimensional spectra processed with nmrPipe (*30*) and analyzed with the program nmrView (*31*). Both {^{1}H, ^{15}N}-HSQC spectra (*32*) and {^{1}H, ^{13}C}-HSQC spectra (*33*) were acquired with 16 transients and 128 increments, zero filled and the number of points doubled with linear prediction during spectral processing.

RESULTS

Figure II-4 shows the amino acid sequence of each peptide cloned into the GB1 fusion protein, the structure of the 56 amino acid globular GB1 protein and the primary sequence of the protein. In this study, I have used peptides of various lengths (12 to 26), hydrophobicities and predicted secondary structure to evaluate the efficiency of peptide production. The cloning site for the inserted synthetic peptide oligonucleotides is C-terminal to the GB1 domain and N-terminal to a poly-6-his tag. Three methionine residues are present in each fusion construct; one on the N-terminus to initiate transcription (start codon) which is not post-translationally cleaved, and two met residues immediately N- and C-terminal to the peptide of interest which were coded for within the synthetic DNA oligonucleotides.

MTYKLILNGK TLKGETTTEA VDAATAEKVF KQYANDNGVD GEWTYDDATK TFTVTE



Peptide	Sequence
cIp-RR-20	TQKIFDLRGKFKRPTLRRVR
cIp-RG-20	TQKIFDLRGKFKRPTLRGVR
cIp-GR-20	TQKIFDLRGKFKRPTLGRVR
TRTK-12	TRTKIDWNKILS
EDQL-26	EDQLREKAKELWQSIYNLEAEKFDLQ
FLQS-26	FLQSDVFFLFLLPPIILDAGYFLPLR

Figure II-4. **GB1 fusion protein constructs**. The primary amino acid sequence of the 56-residue GB1 (2GB1.pdb) protein is shown above with secondary structural details depicted with green arrows (β-sheet) and red cylinders (α-helix). The solved structure of the GB1 domain is shown to the bottom left (*16*). GB1 fusion protein constructs are shown to the right, revealing the location of Methionine residues (denoted by M), with fusion peptide names and amino acid sequences.

Figure II-5 shows the SDS-PAGE gel of the expression of the 6 fusion protein constructs. Lane 1 contains a low molecular weight (MW) marker (biorad), with MW standards as indicated. For simplicity, I have chosen to show the uninduced and induced lanes (lanes 2, 3) for the GB1_cIp-RR-20_His construct, followed by French press lysate supernatant following centrifugation (lane 4) and GB1_cIp-RR-20_His following purification on a Ni^{2+} affinity column (lane 5). As can be observed, the induced GB1_cIp-RR-20_His expressed very well, with an average yield of 80mg/L from expression in 2xYT. I have also included the induced cells and French press lysis supernatant for each of the remaining 5 fusion constructs to compare expression yields and overall solubility of expressed proteins (lanes 6-15). All 6 fusion constructs expressed to approximately equivalent yields, with uninduced lanes being identical (not shown). As can be seen from the gel, all 6 fusion protein constructs produce a soluble fusion protein following French pressure lysis, regardless of amino acid composition of the desired peptide. This is of particular importance regarding construct GB1_FLQS-26_His, as the peptide insert of FLQS-26 is a predicted transmembrane helix with a majority of hydrophobic residues. Thus the ability of the GB1 domain to solubilize a fusion protein-peptide construct is evidenced, as well all 6 constructs are well behaved on a Ni^{2+} purification column producing pure fusion protein constructs. The construct GB1_FLQS-26_His unexpected runs at a lower molecular weight (lanes 14-15) than would be expected when compared with fusion construct GB1_EDQL-26_His (lanes 12-13), which shares an approximate equivalent molecular weight. This is attributed to the difference in amino acid sequence with an increase in hydrophobic amino acids for the GB1_FLQS-26_His construct.

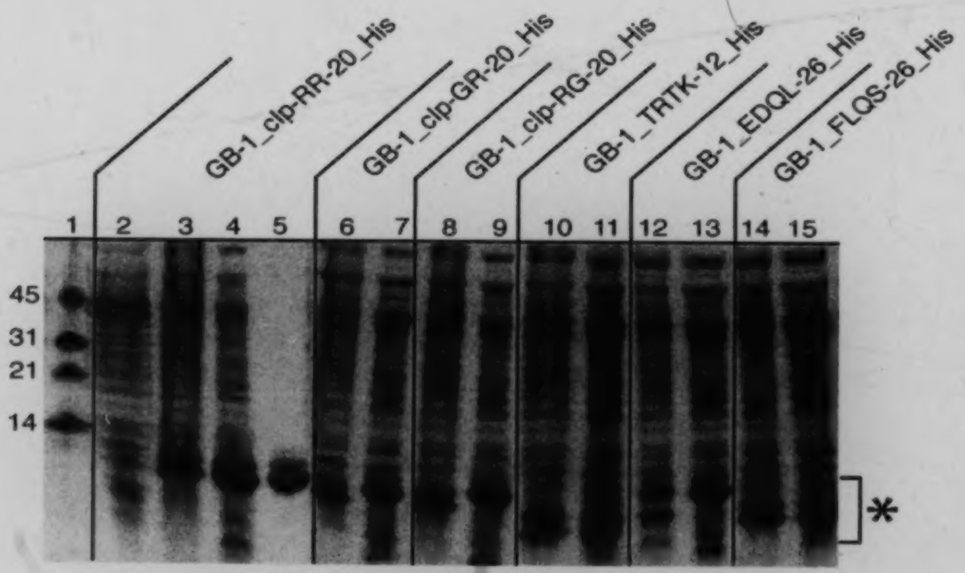


Figure II-5. **8-22% gradient SDS-PAGE gel of expressed fusion proteins**. MW markers (lane 1), uninduced cells (lane 2), induced cells (lane 3), French press lysate supernatant (lane 4) and purified GB-1_clp-RR-20_His (lane 5). Lanes 6-15 contains induced cells and French press lysate supernatant of each indicated fusion protein construct, with regions of fusion protein constructs denoted by *.

Cleavage of the GB1 fusion constructs was performed by CNBr cleavage, which under acidic conditions (0.1N HCl) will cleave C-terminal to all methionine residues, chemically modifying all methionine residues into homoserine residues (see Figure II-6). Monitoring of ε-CH_3 groups of the fusion protein GB1_RR-20_His by ^{1}H NMR spectroscopy (see Figure II-7) indicates that cleavage is quantitative after 12 hours at room temperature as the peak intensity of the ε-CH_3 groups diminishes over time. ε-CH_3 groups of methionine residues will resonate as a sharp singlet within the ^{1}H-NMR spectrum around 2.0ppm, as no J couplings of the methyl group protons to neighboring side chain CH_2 groups are observed due to a separation by a sulfur atom within the side chain. Three singlets are expected for GB1_RR-20_His as three methionine residues are present within the construct, yet only two peaks observed as there is peak overlap of two ε-CH_3 groups at 2.1ppm. The amide bond linkage between Met-Thr residues has been previously reported to be inefficient under standard CNBr cleavage conditions, yet I have proven greater than 99% cleavage for fusion constructs GB1_clp-RR-20_His, GB1_clp-RG-20_His, GB1_clp-GR-20_His and GB1_TRTK-12_His, all of which contain a Met-Thr cleavage site. A review of the literature reveals that I was the first (as far as I can tell) and only person to ever monitor and publish a CNBr cleavage via ^{1}H-NMR.

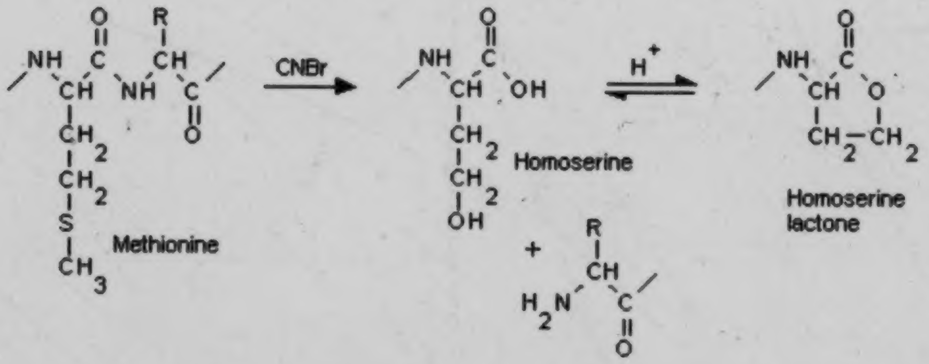


Figure II-6. **Cyanogen Bromide cleavage of amide bond**. The destruction of the amide bond within the protein is irreversibly destroyed C-terminal to all Met residues, leaving either a C-terminal homoserine and/or homoserine lactone.

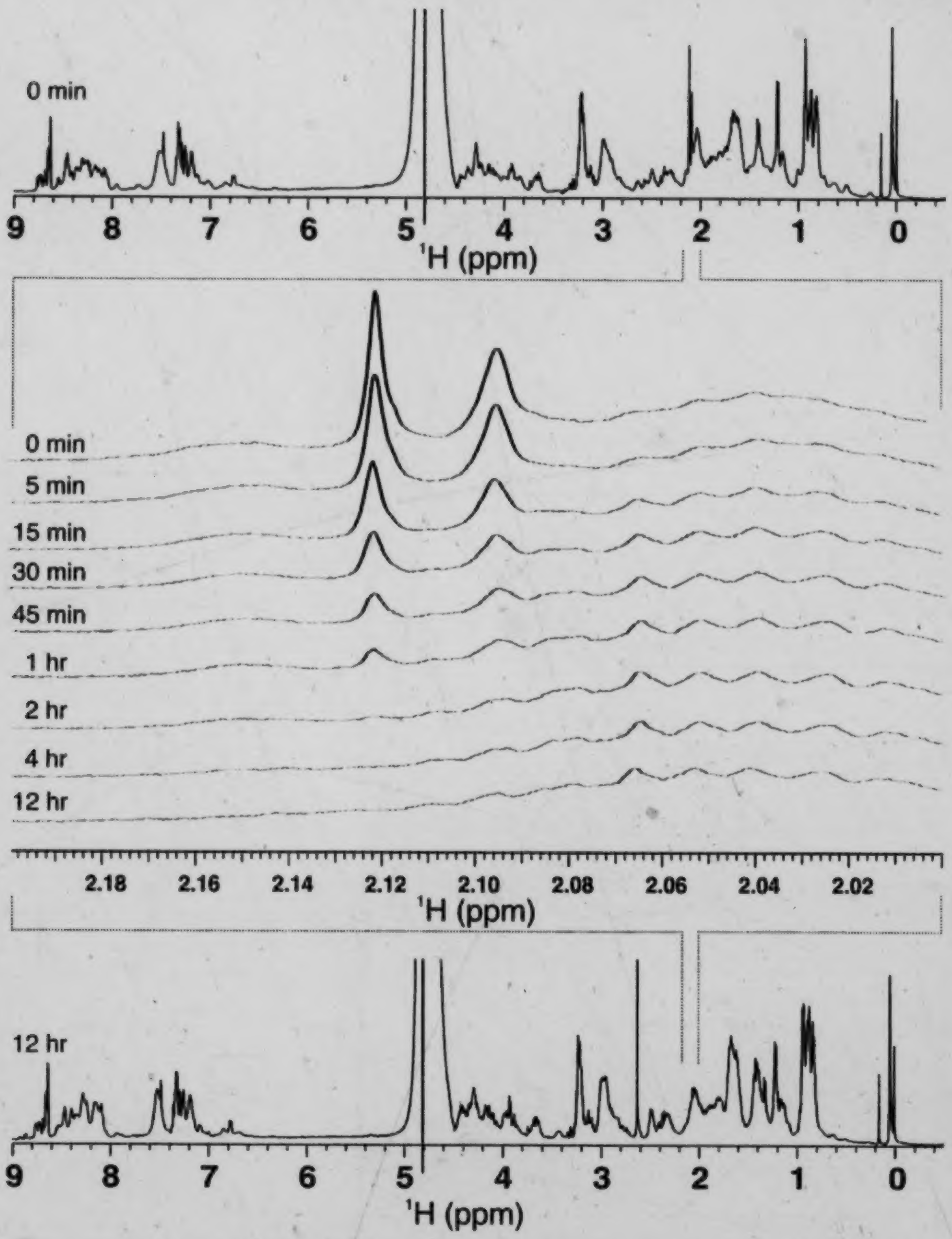


Figure II-7. **Cyanogen Bromide cleavage of GB1 fusion protein.** CNBr cleavage of GB-1_clp-RR-20_His fusion protein is quantitative over 12 hours as monitored by ^{1}H-NMR. Reduction of peak intensity of methionine ε-CH_3 side chains as a function of time indicates efficient cleavage of peptide amide bond.

Following cleavage by CNBr, HPLC purification was performed and verification of protein peaks was obtained by MALDI-TOF mass spectrometry. For simplicity of this thesis chapter, I will only report the results for the fusion protein GB1_RR-20_His, although all fusion proteins presented similar HPLC and mass spectrometry (MS). The HPLC chromatogram can be seen in Figure II-8, with the cleaved fusion protein yielding three domains. The first domain to elute was the His-tag (predicted MW of 1547.7 Da) and was verified by MS to have a molecular weight of 1547.6 Da. The second domain to elute was the peptide clp-RR-20 containing two peaks due to the presence of homoserine and homoserine lactone functional groups, which have predicted MW's of 2618.2 and 2600.2 Da respectively and were verified by MS to be 2617.5 and 2599.5 Da. The third domain to elute was GB1 also contained two main peaks with predicted masses of 6606.4 and 6624.4 Da respectively and were verified by MS to be 6599.4 and 6616.8 Da. The HPLC and MS results for all cleaved peptide fractions for GB1_RR_20_His are detailed in Appendix A-4.

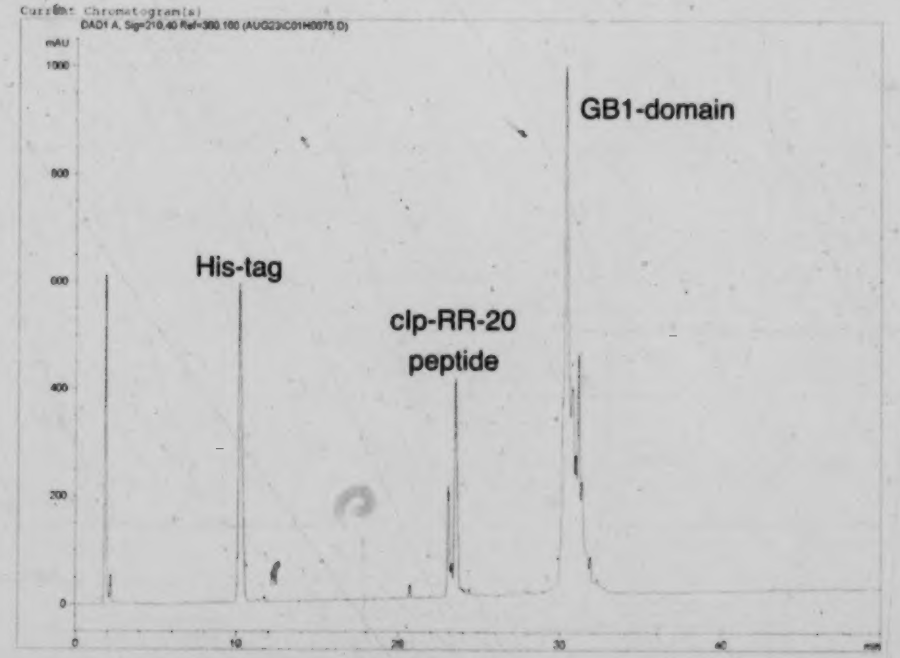


Figure II-8. **HPLC chromatogram of CNBr cleaved fusion protein.** CNBr cleavage of the fusion protein GB1_RR-20_His yields three peak fractions at elution times of 10, 23, and 31 minutes respectively.

For production of isotopically labeled GB1 fusion protein constructs, expression in M9 minimal media with only ^{15}N-$(NH_4)_2SO_4$ as the sole nitrogen source (for ^{15}N-labeling) or ^{15}N-$(NH_4)_2SO_4$/^{13}C-glucose as the only source nitrogen/carbon source produced very limited yield of fusion proteins. The addition of 50mL of ^{15}N and/or $^{15}N/^{13}C$ enriched media (Cambridge Isotope Laboratories, Inc.) to each liter of minimal media restored expression levels to approximately 75% that of 2xYT media, giving high yields with relatively low cost associated with expression of labeled fusion constructs. The addition of enriched media to the minimal media is predicted to supply minerals and nutrients essential to expression of fusion protein constructs.

In total this GB1 expression system yielded, for all fusion constructs tested and affinity purified, a net return of 80mg of fusion protein per litre by an enriched media such as 2xYT, and a net return of 60mg of fusion protein by an isotopically labeled preparation in M9 minimal media. The resulting recovery of peptide from the CNBr cleavage and the HPLC purification protocols yielded an average of 20mg (2xYT) and 15mg (M9 media) per litre of culture. The FLQS-26 peptide was the only poorly behaved peptide following CNBr cleavage, as once cleaved from the GB1 domain lost solubility in the aqueous buffer, and would precipitate out. Easily dissolved in TFA/acetonitrile buffer prior to HPLC column loading, it did not pose any significant problems to purification protocols.

The efficiency of labeling of clp-RR-20 is exhibited in Figure II-9, with {^{1}H, ^{15}N}-HSQC and {^{1}H, ^{13}C}-HSQC NMR spectra of ^{15}N-clp-RR-20 and $^{15}N/^{13}C$-clp-RR-20 being displayed. The complete backbone amide assignments of clp-RR-20 are shown in Figure II-9A (complete NMR chemical shift assignments for this peptide will be discussed in detail in Chapter III of this thesis); with residues T^{128} and Q^{129} not observed due to exchange with water, and one homoserine resonance is observed. clp-RR-20 is unstructured in solution, as evidenced by the limited spectral dispersion in the {^{1}H, ^{15}N}-HSQC NMR spectrum.

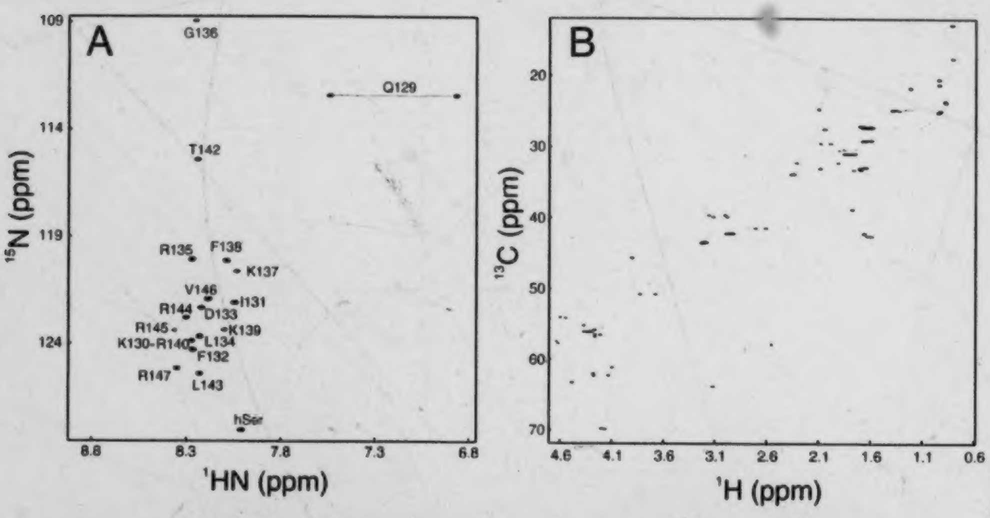


Figure II-9. **HSQC NMR spectra of cIp-RR-20 peptide**. Amide backbone assignments of cIp-RR-20 are shown in the {^{1}H, ^{15}N}-HSQC shown in panel (A), with a single homoserine resonance observed. Panel (B) displays the {^{1}H, ^{13}C}-HSQC NMR spectrum of the peptide.

DISCUSSION

In this chapter, I have demonstrated the practicality of peptide production via a recombinant fusion protein technique using the GB1 domain from *Streptococcus*. More importantly this technique has been shown to be directly applicable to the production of isotopically labeled peptides for use in NMR studies, however this application is not without its limitations. This technique deals exclusively with CNBr cleavage of peptides from the fusion protein, and as such no methionine residues can be present in the sequence of the intended peptides. Should a targeted peptide contain a methionine residue, it might be possible to perform CNBr cleavage if a conserved mutagenesis to another residue is performed, assuming this does not perturb behavior or structure. Should the intended peptide require a methionine residue within its sequence, the potential is available to engineer other cleavage sites into the GB1 fusion protein such as thrombin or other proteases, however this would add unwanted length to the inserted oligonucleotides, which currently lose fidelity in synthesis of over 100 base pairs. The ability to clone beneficial protease cleavage sites directly into the vector would alleviate this problem and could be performed quickly and simply.

When this work was initiated in 2001, there was a severe lack of reported peptide production techniques in the literature. I required an expression system that worked efficiently and did not require large amounts of denaturants, as were the only available options at the time (*13-15*). Since the publication of this work in the August 2003 issue of *Protein Science* (*1*), there has been an increase in competitive procedures that mirror my own in the area of peptide production techniques via recombinant methods. In August of 2003, two weeks following my GB1 publication, Osbourne and co-workers reported robust recombinant peptide production via a fusion with the novel fusion protein SFC120, however this group similarly used the limitation of CNBr cleavage as well for peptide purification and the use of chemical denaturants during purification (*34*). In May of 2004, Tenno and co-workers reported peptide production via a fusion technique with the protein thioredoxin, which produced soluble fusion constructs, followed by peptide cleavage via Factor Xa protease (*35*). This competing method is the only other peptide production technique (in contrast to my own) that produced equivalent amounts of isolated peptides for the intent of NMR studies via a soluble fusion construct.

The ability to produce high amounts of isotopically labeled peptides for use in NMR studies has garnered a large amount of interest from the scientific community. In a collaborative project with Dr. Larry Fliegel from the Department of Biochemistry at the University of Alberta, I produced the peptide FLQS-26 for structural studies of a membrane-spanning region of the human sodium-proton exchanger isoform 1 (hNHE1). The structural data elucidated from this peptide in a lipid-mimetic solvent was a valuable asset to the functional studies of hNHE1 (performed by Ph. D. student Emily Slepkov, Fliegel laboratory), and has cumulated in a publication, which was recently published in May 2005 in the *Journal of Biological Chemistry* [*J Biol Chem.* 2005 May 6 **280(18)**:17863-72]. However, this data is not immediately relevant to the focus of this thesis in regards to cardiac contraction via the thin filament and is therefore provided only as an attached reading in Appendices B and C for reference.

I have demonstrated the feasibility of a GB1 fusion protein approach for use in recombinant peptide labeling and expression, with high purity of product and high product yield. The ability of the GB1 protein to 'protect' attached peptides in bacterial cytoplasm is virtually unique, as opposed to other methods using inclusion body seclusion for protease protection. The applications for labeled peptides for NMR studies is large, yielding the ability to measure relaxation rates of bound peptides to a target protein, as well as giving the ability to perform multi-dimensional heteronuclear NMR experiments for solution structure determination. I have produced the peptide cIp-RR-20 with the sole intention of solving the NMR solution structure of the bound inhibitory region of cTnI to the C-lobe of cTnC for my thesis work (see Chapter III). Given the importance of protein-peptide interactions in cellular functions, it is expected that a well-behaved, well-characterized, and efficient system for expressing and purifying recombinant peptides will be critical for structural and functional studies.

REFERENCES

1. Lindhout, D. A., Thiessen, A., Schieve, D., and Sykes, B. D. (2003) *Protein Sci 12*, 1786-91.
2. Tyrberg, B., and Levine, F. (2001) *Curr Opin Investig Drugs 2*, 1568-74.
3. Hristova, K., Dempsey, C. E., and White, S. H. (2001) *Biophys J 80*, 801-11.
4. Pollesello, P., and Annila, A. (2002) *Biophys J 83*, 484-90.
5. Hutter, M. C., Krebs, J., Meiler, J., Griesinger, C., Carafoli, E., and Helms, V. (2002) *Chembiochem 3*, 1200-8.
6. Mazor, Y., Gilead, S., Benhar, I., and Gazit, E. (2002) *J Mol Biol 322*, 1013-24.
7. Azriel, R., and Gazit, E. (2001) *J Biol Chem 276*, 34156-61.
8. Saiman, L., Tabibi, S., Starner, T. D., San Gabriel, P., Winokur, P. L., Jia, H. P., McCray, P. B., Jr., and Tack, B. F. (2001) *Antimicrob Agents Chemother 45*, 2838-44.
9. Gudmundsson, G. H., Agerberth, B., Odeberg, J., Bergman, T., Olsson, B., and Salcedo, R. (1996) *Eur J Biochem 238*, 325-32.
10. Nathisuwan, S., and Talbert, R. L. (2002) *Pharmacotherapy 22*, 27-42.
11. Jacobsen, J. S. (2002) *Curr Top Med Chem 2*, 343-52.
12. Celis, E. (2002) *J Clin Invest 110*, 1765-8.
13. Majerle, A., Kidric, J., and Jerala, R. (2000) *J Biomol NMR 18*, 145-51.
14. Jones, D. H., Ball, E. H., Sharpe, S., Barber, K. R., and Grant, C. W. (2000) *Biochemistry 39*, 1870-8.
15. Sharon, M., Gorlach, M., Levy, R., Hayek, Y., and Anglister, J. (2002) *Protein Expr Purif 24*, 374-83.
16. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, G. M. (1991) *Science 253*, 657-61.
17. Gronenborn, A. M., and Clore, G. M. (1996) *Prot. Sci. 5*, 174-177.
18. Huth, J. R., Bewley, C. A., Jackson, B. M., Hinnebusch, A. G., Clore, G. M., and Gronenborn, A. M. (1997) *Prot. Sci. 6*, 2359-2364.
19. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.
20. Li, M. X., Spyracopoulos, L., Beier, N., Putkey, J. A., and Sykes, B. D. (2000) *Biochemistry 39*, 8782-8790.
21. Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) *J Mol Biol 271*, 728-50.
22. Campbell, A. P., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) *Biochim Biophys Acta 1160*, 35-54.
23. Hernandez, O. M., Housmans, P. R., and Potter, J. D. (2001) *J Appl Physiol 90*, 1125-36.
24. Redwood, C. S., Moolman-Smook, J. C., and Watkins, H. (1999) *Cardiovasc Res 44*, 20-36.
25. Ivanenkov, V. V., Jamieson, G. A., Jr., Gruenstein, E., and Dimlich, R. V. (1995) *J Biol Chem 270*, 14651-8.
26. Counillon, L., Noel, J., Reithmeier, R. A., and Pouyssegur, J. (1997) *Biochemistry 36*, 2951-9.
27. Wakabayashi, S., Pang, T., Su, X., and Shigekawa, M. (2000) *J Biol Chem 275*, 7942-9.

28. Pearlstone, J. R., and Smillie, L. B. (1985) *Can J Biochem Cell Biol 63*, 212-8.
29. NCBI. (2005) *National Center for Biotechnology Information*, http://www.molbiol.ox.ac.uk/~cocallag/refdata_html/codonusagetable.shtml.
30. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
31. Johnson, A. J., and Blevins, R. A. (1994) *Journal of Biomolecular NMR 4*, 603-614.
32. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.
33. Neri, D., Szyperski, T., Ottig, G., Senn, H., and Wuthrich, K. (1989) *Biochemistry 28*, 7510-7516.
34. Osborne, M. J., Su, Z., Sridaran, V., and Ni, F. (2003) *J Biomol NMR 26*, 317-26.
35. Tenno, T., Goda, N., Tateishi, Y., Tochio, H., Mishima, M., Hayashi, H., Shirakawa, M., and Hiroaki, H. (2004) *Protein Eng Des Sel 17*, 305-14.

CHAPTER III

The inhibitory region of cardiac troponin I: structure and function

OVERVIEW

This chapter presents NMR structural studies of the inhibitory region of cardiac troponin I, in complex with the C-lobe of cardiac troponin C in the calcium saturated state. In addition to the solution structure of this complex, I will present ^{15}N-NMR backbone amide relaxation data that aids in the understanding of the interactions occurring between the two domains. The majority of the data presented in this chapter was published in the *Journal of Biological Chemistry* on July 18th, 2003 [J Biol Chem. 2003 Jul 18;**278(29)**:27024-34]. The peptide cIp-RR-20, whose expression and purification was previously described in Chapter II, shall be referred to as cIp (cardiac inhibitory peptide – wild-type) from this point onwards.

Cardiac troponin C is the Ca^{2+}-dependent switch for heart muscle contraction. Troponin C is associated with various other proteins including troponin I and troponin T (*1-9*). The interaction between the subunits within the troponin complex is of critical importance in understanding contractility. Following a Ca^{2+} signal to begin contraction, the inhibitory region of troponin I comprising residues Thr128-Arg147 relocates from its binding surface on actin to troponin C, triggering movement of troponin-tropomyosin within the thin filament and thereby freeing actin binding site(s) for interactions with the myosin ATPase of the thick filament to generate the power stroke (*7, 10, 11*). The structure of calcium saturated cardiac troponin C (C-lobe) in complex with the inhibitory region of troponin I was determined using multinuclear and multidimensional nuclear magnetic resonance spectroscopy. The structure of this complex reveals that the inhibitory region adopts a helical conformation spanning residues Leu134-Lys139, with a novel orientation between the E and H helices of troponin C, which is largely stabilized by electrostatic interactions. Using isotope labeling, I have studied the dynamics of the protein and peptide in the binary complex. The structure of this inhibited complex provides a framework for understanding interactions within the troponin complex upon heart contraction.

INTRODUCTION

TnI-TnC interactions have been studied extensively for over 3 decades. The functionality of domains of TnC was determined early by Head and Perry (*12*) using various proteolytic/cleavage techniques. These studies of rabbit skeletal muscle yielded a region of TnI which has been referred to as the 'inhibitory region'. Subsequent work by Talbot and co-workers identified a refined inhibitory region, which has been mapped to a central region of its primary amino acid sequence (*13*), corresponding to residues Thr128-Arg147 in the cardiac system (cIp). The inhibitory region contains the minimum sequence required to fully inhibit the myosin ATPase activity on the thin filament (*14*). Interactions of the cardiac TnI inhibitory region are of key importance in complete understanding of muscle regulation. It has been shown that the inhibitory region interacts with the protein actin on the thin filament during muscle relaxation, and that movement of cIp off actin to interact with TnC occurs upon Ca^{2+}-induced muscle contraction. Previously, it has been shown that the inhibitory region of TnI binds to the C-lobe of TnC and the central linker region of the N- and C-terminal domains of TnC (*15*). Additional interactions of TnC with other regions of TnI have been elucidated (*16, 17*), resulting in an intertwined system with multiple contacts between the subunits of the troponin complex.

The wealth of knowledge obtained by functional studies of cardiac (and skeletal) contracting tissues over many decades has led to a *simple*, yet *difficult* question among biophysicists: "What, *specifically*, is occurring at the *molecular level* during muscle contraction?". Atomic resolution structural explanation of the protein-protein interactions involved in contraction would be an invaluable asset to the current understanding within the field of muscle physiology. Specifically, the understanding of the calcium-signaling cascade within the troponin complex, and domain (re)-orientations would be a welcomed addition to which years of biophysical data could be measured against. The *difficulty* in pursuing structural studies of the troponin complex lies in the flexibility of the domains within this ternary protein complex, and solubility of the troponin I and T subunits. Hence, structural studies relying solely on crystallography attempts have met with limited success do to this intrinsic mobility. When my thesis project was started in February 2000, there was only *one* published crystal structure involving troponin I, in which an N-terminal fragment (RP40) was bound to the C-lobe of troponin C (*18*). The globular nature of TnC makes it a prime candidate for crystallization studies; however, the flexible nature of the troponin I and T makes it difficult (if not impossible) for crystallization of the larger, associated complex. Hence structural studies are more suited to solution-based experimentation such as NMR, as it can overcome intrinsic areas of flexibility for structural elucidation. However, due to molecular size limitations and solubility limitations of individual subunits, NMR has been limited in working with smaller fragments/domains of the troponin I and T subunit(s) (as defined by Perry and Head (*12*)), in association with the globular TnC subunit(s). NMR experimentation can therefore succeed where crystallography has met limited success in the past. In addition to structural elucidation, NMR has an added ability to directly address potential domain flexibilities via ^{15}N- and/or ^{13}C-relaxation of backbone amide and side-chain methyl groups. Thus NMR is the choice to pursue structural studies of the troponin complex.

When I began my graduate career in the laboratory of Dr. Brian D. Sykes in early 2000, I was challenged to solve the structure of the critical inhibitory region in complex with troponin C. Prior to my arrival, NMR chemical shift mapping experimentation by Dr. Monica Li in the Sykes laboratory had shown that it binds to TnC on/near the C-lobe and central D/E helix (*15, 19*). Elucidation of the structure proved to be much more difficult and time intensive than I had estimated, as production of labeled inhibitory peptide was required for unambiguous assignment of inter-domain NOEs for final structural calculation. The production strategy for labeled peptide is discussed in detail in Chapter II of this thesis. Structural determination of the inhibitory region in complex with cTnC in the calcium-saturated state would be a critical structure regarding the placement of this region within the troponin complex during calcium-induced muscle contraction. More specifically, I was looking to address individual interactions that are critical for complex formation, which may become only available in the calcium saturated (contracting) state. Could I really begin to understand why this inhibitory region is so critical to the whole of muscle contraction, and why it is so well conserved throughout species and tissues? Could a connection be established as to why this region contains a regulatory Protein Kinase C (PKC) site on Thr142 and how this directly affects muscle contraction (*20, 21*)? More specifically, why does only the cardiac isoform contain a Thr142 when other isoforms such as skeletal contain a Pro in place (*22-26*)? As well, how does a Familial Hypertrophic Cardiomyopathy (FHC) mutation R144G affect muscular contraction (*27*)? Can a direct correlation to physiological symptoms be established, and in turn, could I predict the consequences of other potentially deleterious mutations on this region? It is due to the pursuit in attempting to answer these questions that I feel an atomic-resolution structure of this region in complex with TnC must be solved.

I am not the first person to directly address structural studies of the inhibitory region. I was therefore able to 'hit the ground running' as I could learn and build upon the results of those who came before me. The one lesson that can be learned from all the previous groups who have researched the inhibitory region, in both the cardiac and skeletal isoforms, is that there is an *agreement not to agree*. It would be pointless for myself to begin structural studies of the inhibitory region if it had already previously been solved. I therefore set out to perform two duties: to succeed where others have failed and solve the structure of the inhibitory region bound to troponin C *and* to try to put a finality on the disagreements which have surrounded this region.

Due to the critical importance that it plays in muscle contraction, it is not at all surprising that many groups have intensely studied the inhibitory region (*18, 19, 28-38*). However, despite all the combined efforts, the structure has remained elusive. Thus there have been many proposed models for the placement and structure of this region, some of which are directly based upon biophysical results. Early ^{1}H-NMR studies by Campbell involving transferred NOEs using a small, 11-residue peptide of the skeletal isoform ($sTnI_{104-115}$), which correspond to $cTnI_{136-147}$ in the cardiac isoform, led to the proposal that the inhibitory region adopts a short helix, distorted around two central proline residues (*30, 31*). Subsequent docking of this structure with sTnC placed this distorted helix within the hydrophobic cleft on the C-lobe (*33*). Using a paramagnetic relaxation reagent, Howarth proved that cIp binding to cTnC protects the surface exposure rates of Met 120 and 157, but not Met 103 or Met137 of the C-lobe (*29*). This data was in agreement with a binding region in the vicinity of the hydrophobic face of the C-lobe. A crystal structure of sTnC in complex with the N-terminal regulatory peptide of $sTnI_{1-47}$ (sRP40) presented by Vassylyev and coworkers (*18*) showed that this region adopts a helical structure and binds in the hydrophobic cleft of sCTnC. This structure directly conflicted with the Campbell result as superimposition of the two structures revealed a steric clash between the TnI domains. To deal with the conflict, Vassylyev modeled the inhibitory region away from the hydrophobic cleft, but kept with the premise that it adopts a helical conformation. However, the conformation of the inhibitory region was directly challenged by Hernandez and coworkers (*28*), who via ^{1}H-NMR chemical shift data proposed an extended conformation of the inhibitory region, with a two-stranded β-hairpin away from the hydrophobic cleft of the C-lobe (*38*). Work on the skeletal isoform by former fellow graduate student Dr. Pascal Mercier in the Sykes laboratory, revealed that the RP40 region will displace the inhibitory region off the C-lobe, and proposed that the inhibitory region will shift its position on the C-lobe upon RP40 binding and bind over RP40 across the hydrophobic face of the C-lobe (*19, 32*).

Electron spin labeling work by Brown with the cardiac isoform confirmed that the N-terminal section of the inhibitory region ($cTnI_{129-137}$) displays a helical profile, with the C-terminal residues 139-145 displaying no discernable secondary structural characteristics (*34*). The inhibitory region possesses a large number of basic residues, and Tripet (*35*) predicted that this highly basic region of cTnI makes numerous electrostatic interactions with the acidic TnC in the troponin complex. The majority of these basic residues are within the C-terminal tail of the inhibitory region, which might agree with an extended conformation as proposed by Brown. In addition, recent work on the cardiac isoform using residual dipolar couplings (RDCs) by Dvoretsky and coworkers determined the orientations of the domains of cTnC within a cTnI•cTnC complex (*36*), and a small angle scattering study by Heller and coworkers has determined relative domain orientation within a cTnI•cTnC•cTnT$_{198\text{-}298}$ complex (*37*) which suggests that interactions between TnC and the TnI•TnT components differs significantly between the skeletal and cardiac isoforms.

With the growing amount of data regarding the inhibitory region and its role in muscular contraction/regulation, it was a prime time for myself to elucidate the structure of the bound inhibitory region and finally lay to rest a problem that has plagued the field since the time of Perry and Head: what does the inhibitory region look like when bound and how are these interactions occurring?

In this chapter, I will discuss my results regarding the successful elucidation of the NMR solution structure of the cardiac isoform of the inhibitory region of TnI (cTnI$_{128\text{-}147}$) in complex with the Ca^{2+} saturated C-terminal domain of TnC. I chose to solve the bound structure to only the C-lobe of cTnC to reduce the size of the complex to ease in spectral assignment, as previous chemical shift mapping results indicated that it only binds to the C-lobe (*15*). In the bound structure, the inhibitory region displays a helical secondary structure from residues Leu134-Lys139, with several stabilizing electrostatic interactions with cCTnC. This binary structure correlates well with previous NMR chemical shift mapping of the interactions of the inhibitory region with TnC (*15, 19, 27*) and with the electron spin labeling results by Brown (*34*). Furthermore, with the ability to isotopically label the inhibitory region, I was able to utilize ^{15}N-NMR relaxation to study dynamics of the bound inhibitory region. NMR relaxation indicates that the central core of cIp is rigid when bound to cCTnC. Unexpectedly the C-terminal tail of the inhibitory region (Arg140-Arg147), which displays no discernable secondary structural elements or hydrophobic NOEs to the C-lobe, is rigidly bound to the C-lobe. This is the first high-resolution structure determined of the inhibitory region of TnI in complex with TnC, and provides a framework for understanding interactions within the troponin complex during heart contraction.

On July 3rd, 2003, a publication in the journal *Nature* by Soichi Takeda out of the RIKEN structural biology institute in Japan (Maeda laboratory), revealed simultaneous 2.6/3.3Å crystal structures of 46/52kDa cardiac troponin complexes which consisted of ternary complexes of $cTnC{\bullet}cTnI_{31\text{-}210}{\bullet}cTnT_{183\text{-}288}$ in the calcium saturated state (*39*). This was a welcomed addition to the field of muscle structural biology as we finally had definitive insights into muscle contraction in the form of an intact troponin structure. These high-resolution structures were useful in visualizing many of the key troponin interactions at the atomic level, within the crystal lattice. However these structures were incomplete as electron densities were missing in key areas such as the inhibitory region, results which again re-enforced the emphasis that there is a large degree of flexibility occurring within the troponin complex. This structure confirmed the hypothesized coiled-coil interactions occurring between troponin I and T subunits; however, the head orientations of the N- and C-lobes of cTnC were not as were predicted by ^{15}N-RDC studies by NMR (*36*). Therefore, there is still much discussion in the field of muscle structural biochemistry, as with as many questions as this crystal structure answered, it yielded the same number of new questions. It was therefore entirely relevant and timely that my published NMR solution structure of the inhibitory region bound to the C-lobe was within one week of this crystal structure as it can address aspects of the crystal structure that were unclear. To date, my binary structure of the bound inhibitory region remains the only atomic resolution structure of the key inhibitory region for the cardiac isoform.

EXPERIMENTAL PROCEDURES

III-A. Preparation of cCTnC protein

The engineering of the expression vector for the cCTnC (90-161) protein was as described by Chandra (*24*). For the expression and purification of cCTnC, *E. coli* BL21 (DE3) pLysS cells were transformed (heat shock – refer to Chapter II experimental procedures) with a pET-3a vector containing the cCTnC clone. 100mL starter cultures of 2xYT/Amp were inoculated with the transformed cells and left overnight at 37°C with shaking. One litre cultures of 2xYT were inoculated with 10mL of the overnights and left to grow at 37°C with shaking until an OD_{600} of 0.8 was achieved. IPTG was then added to a final concentration of 1mM for induction and left at 37°C with shaking for a further 4 hours. Cells were then pelleted by centrifugation (2 minutes, 6K rpm), and re-suspended in 25mL (per 1L of culture) of 50mM Tris (pH 8.0) containing 10nM PMSF and frozen at -20°C overnight. Following thawing, 50*u*L of DNase I (18mg/mL) was added with 50mg of $MgSO_4$ to the cell suspension and left on ice for 1 hour. The cell suspension was passed through a French Pressure Cell 3 times at 1000psi for cellular lysis and the supernatant following centrifugation (60 minutes, 16K rpm), was added directly to a DEAE anion exchange column equilibrated with column buffer (50mM Tris, 0.1M KCl, pH 8.0). A linear gradient was established (0.1M KCl → 0.4M KCl) and fractions were monitored by SDS-PAGE (*40*). Fractions containing the C-lobe were subjected to a G-25 sephadex de-salting column (10mM NH_4HCO_3, pH 8.0) and lyophilized to dryness. A second round of DEAE/sephadex G-25 column purification was performed to increase purity of cCTnC. Semi-pure cCTnC was further purified on a Superdex 75 (Amersham Pharmacia) column (50mM Tris, 150mM NaCl, pH 8.0) and desalted on a G-25 de-salt column, followed by lyophilization to dryness (*27, 41*).

III-B. Preparation of cIp protein

Synthetic cIp peptide acetyl-`TQKIFDLRGKFKRPTLRRVR`-amide was prepared as described for the sIp peptide in Tripet (*17*). The engineering of the expression vector for expression and purification of ^{15}N-cIp and $^{13}C/^{15}N$-cIp via a fusion protein approach has been previously described in Chapter II of this thesis (*27, 42, 43*).

III-C. Isotope labeling of cCTnC and cIp proteins

For isotope incorporation (i.e. ^{15}N and/or ^{13}C) into the cCTnC and/or cIp proteins, induction cultures consisted of M9 minimal media (50mM Na_2HPO_4, 25mM KH_2PO_4, 10mM NaCl, pH 7.3) containing 2.5g ^{13}C-glucose and/or 1.1g ^{15}N-$(NH_4)_2SO_4$ per litre. Induction and purification procedures are identical to those of unlabelled protein.

III-D. ^{1}H-NMR structural studies on unbound clp protein

6.41mg of synthetic clp peptide was dissolved in 600*u*L of NMR buffer (100mM KCl, 10mM Imidazole, 90%H_2O/10%D_2O, pH 6.7) to which was added 10*u*L of 10mM DSS, 10*u*L of 1.5% NaN_3, and 5*u*L of 1M $CaCl_2$, and 500*u*L was added directly to an NMR tube. All NMR experiments and parameters for ^{1}H-clp are reported in *Table III-1*.

III-E. Titration of synthetic clp peptide with ^{15}N/^{13}C-cCTnC•2Ca^{2+}

6.32mg of synthetic clp peptide was dissolved in 600*u*L of NMR buffer (100mM KCl, 10mM Imidazole, 90%H_2O/10%D_2O, pH 6.7) to which was added 10*u*L of 10mM DSS, 10*u*L of 1.5% NaN_3, and 5*u*L of 1M $CaCl_2$, and 500*u*L was added directly to an NMR tube. Solid, lyophilized ^{15}N/^{13}C -cCTnC was added directly to the NMR tube in additions of 0.43, 0.50, 1.19, 1.43, 1.48, 2.12, 2.19, 2.18, 2.30 and 5.76 mg quantities. Changes in ^{1}H-clp chemical shifts upon binding were monitored via ^{15}N/^{13}C-filtered DISPI NMR spectroscopy at 500MHz.

III-F. Titration of ^{15}N-cCTnC•2Ca^{2+} with synthetic clp peptide

To an NMR tube containing a 500*u*L sample of 0.59 mM ^{15}N-cCTnC•2Ca^{2+} were added aliquots of 1.8*u*L of 45.1 mM synthetic-clp stock solution for the first 5 titration points. Aliquots of 1.9*u*L were added for the next 7 titration points, followed by aliquots of 4.9*u*L for the next 3 titration points, followed by a final addition of 9*u*L to ensure complete clp saturation. The sample was thoroughly mixed after each addition. After every titration point, 1*u*L of the resulting titrated solution was removed and used for amino acid analysis. The change in cCTnC•2Ca^{2+} concentration due to increased volume during the titration was taken into account for data analysis, the change in pH for clp addition was negligible. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC spectra were acquired at every titration point. In order to check if the synthetic and recombinant clp peptides bound in an identical fashion to cCTnC•2Ca^{2+}, 8.1 mgs of recombinant ^{1}H-clp was added to a 0.92 mM ^{15}N-cCTnC•2Ca^{2+} in NMR buffer containing 5*u*L of 1M $CaCl_2$ to generate a complex in which the $[\text{clp}]_{total}/[\text{cCTnC}\bullet2\text{Ca}^{2+}]_{total}$ was approximately 5:1, corresponding to the final ratio for the synthetic complex. 2D {^{1}H, ^{15}N}-HSQC spectra were acquired for each sample and superimposed to observe any potential differences on the cCTnC backbone when the two peptides were bound.

III-G. Titration of ^{15}N-clp with cCTnC•2Ca^{2+}

5.11mg of recombinant ^{15}N-clp was dissolved in 550*u*L of NMR buffer (pH 6.7) containing 25*u*L 1M $CaCl_2$, and 500*u*L was transferred to an NMR tube. Solid cCTnC was added in 1.2mg additions, with thorough mixing after each addition. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC spectra were acquired at every titration point. After every titration point, 1*u*L of the resulting titrated solution was removed and used for amino acid analysis. Solid cCTnC was added until no further changes were observed in the HSQC spectra, ensuring complete clp saturation. The change in clp concentration due to changes in volume during the titration was taken into account for data analysis, the change in pH for cCTnC addition was corrected by addition of 1M NaOH during each addition. A small amount of white precipitate accumulated as increasing amounts of solid cCTnC was added.

III-H. Structural studies on cCTnC•2Ca^{2+}•clp complex

Three samples were prepared for structural studies on the complex, with subsequent structural determination carried out using the NMR experiments listed in *Table III-II*. 10.13mg of ^{13}C/^{15}N-cCTnC was dissolved in 600*u*L NMR buffer (pH 6.7) containing 5uL of 1M $CaCl_2$ to which 6.23mg of synthetic clp was added, and 500*u*L was added to an NMR tube. 3.34mg of ^{13}C/^{15}N-clp was dissolved in 600*u*L of NMR buffer (pH 6.7) containing 5*u*L of 1M $CaCl_2$ to which 25.5mg of cCTnC was added to ensure complete clp saturation, and 500*u*L was added to an NMR tube. 5.06mg of ^{13}C/^{15}N-clp was dissolved with 16.44mg of ^{13}C/^{15}N-cCTnC to ensure a 1:1 molar ratio in 575*u*L of NMR buffer (100% D_2O, pH 6.7) containing 5*u*L of 1M $CaCl_2$ and 500*u*L was added to an NMR tube.

III-I. NMR Spectroscopy

All NMR data used in this study were acquired at 30°C using Varian INOVA 500 MHz, Unity 600 MHz, and INOVA 800 MHz spectrometers. All three spectrometers are equipped with triple resonance probes and Z-axis pulsed field gradients (XYZ gradients for the 800 MHz). For the unbound cIp peptide and the cCTnC•cIp complex, the chemical shift assignments of the backbone and the side chain atoms and NOE interproton distance restraints were determined using the two-dimensional and three-dimensional NMR experiments described in *Tables III-I and III-II*.

III-J. Data processing and peak calibration

All two-dimensional and three-dimensional NMR data was processed using NMRPipe (*44*), and all one-dimensional NMR data was processed using VNMR (Varian Associates). All spectra were analyzed using nmrView (*45*). For cCTnC and cIp in the complex, intramolecular distance restraints obtained from the NOESY experiments were calibrated according to Gagne *et al.* (*46*). Dihedral angle restraints were derived from data obtained from HNHA, and NOESY-HSQC experiments according to Sia (*5*).

III-K. Structural calculations for binary complex cCTnC•2Ca^{2+}•cIp

NOE contacts between protein-protein, peptide-peptide and protein-peptide were derived from various 3D-^{13}C/^{15}N-noesy experiments. The 3D-^{13}C-edited noesy experiment performed on ^{13}C/^{15}N-cCTnC complexed with ^{13}C/^{15}N-cIp was used to define protein-peptide contacts with use of symmetrical peaks within the 3D planes for use in unambiguous NOE assignments. All protein-peptide NOEs were calibrated to 4 ± 2 Å. An initial set of 100 structures of both cCTnC and cIp were first generated separately using NOE distance restraints only. NOEs with a distance violation of 0.2 Å or greater were closely examined prior to further rounds of structure refinements. In the latest stages of refinement, ϕ angles of cCTnC were added for residues located in well-defined regions, as determined using the program Procheck (*47*). In the first round of structure calculations of the cCTnC-cIp complex, an initial set of 100 structures was determined using pre-folded structures of cCTnC and cIp respectively, with unambiguous protein-peptide NOEs being introduced. Further refinement of structures of the complex using the initial set of 100 structures was performed using the program procheck to inspect structures, as well as closely examining NOE distance violations greater than 0.1 Å. To ensure independent folding of the complex from pre-folded cCTnC and cIp structures, a final set of 100 structures was derived from both cCTnC and cIp starting in extended conformations. The default CNS (*48*) built-in annealing protocol was modified to allow the introduction of 12 Ca^{2+}-distance restraints preceding the second cooling step using Cartesian dynamics. The structural statistics of the family of the 30 calculated lowest energy structures is shown in *Table III-III*. A comparison of interhelical angles of the binary complex with previously solved EF hand structures is shown in Table *III-IV*.

III-L. Backbone amide ^{15}N relaxation measurements of cCTnC•2Ca^{2+}•cIp complex

Two samples of ^{15}N-cCTnC were prepared, one of cCTnC•2Ca^{2+} free in solution and the other with binary complex cCTnC•2Ca^{2+}•cIp. A first sample of 11.82mg of ^{15}N-cCTnC was dissolved in 600*u*L of NMR buffer (pH 6.7) containing 5*u*L of 1M $CaCl_2$ and 500*u*L was added to an NMR tube. A second sample of 12.49mg ^{15}N-cCTnC was dissolved in 600*u*L NMR buffer (pH 6.7) containing 5*u*L of 1M $CaCl_2$ to which 6.69 mg of cIp peptide was added to ensure complete cCTnC saturation and 500*u*L was added to an NMR tube. Two samples of ^{15}N-cIp were prepared, one of cIp free in solution and the other with cIp•cCTnC•2Ca^{2+} in complex. 0.8mg of recombinant ^{15}N-cIp was dissolved in 550*u*L of NMR buffer (pH 6.7) containing 5*u*L of 1M $CaCl_2$ and 500*u*L was added to an NMR tube. A second sample of 1.88mg of recombinant ^{15}N-cIp was dissolved 600*u*L of NMR buffer (pH 6.7) containing 5*u*L of 1M $CaCl_2$ to which 14.5mg of ^{1}H-cCTnC was dissolved to ensure complete ^{15}N-cIp saturation and 500*u*L was added to an NMR tube.

All relaxation data were acquired on Varian INOVA 500 MHz and Unity 600 MHz spectrometers at 30°C. Sensitivity enhanced pulse sequences developed by Farrow et al. (*49*) were used to measure ^{15}N-T_1, ^{15}N-T_2, and $\{^1H\}$-^{15}N NOE. Using two-dimensional spectroscopy a set of backbone ^{15}N-T_1, ^{15}N-T_2, and {^{1}H}-^{15}N NOE experiments were collected for each sample with parameters described in *Table III-V*. The delay between repetitions of the pulse sequence was set to 3s for both the T_1 and T_2 experiments. {^{1}H}-^{15}N NOE measurements (η = [NOE-noNOE]/noNOE) were made in the absence (relaxation delay incorporation of 5s between spectrometer pulses) and the presence or proton saturation (incorporation of 3s of ^{1}H saturation, with a delay of 2s between spectrometer pulses). All relaxation data were processed using NMRPipe (*44*) and analyzed using nmrView (*45*).

III-M. Coordinates

The coordinates for the structure of cCTnC•2Ca^{2+}•cIp have been deposited in the RCSB Protein Data Bank with the assigned accession code ***1OZS***.

RESULTS

III-A. Structural studies of unbound synthetic ^{1}H-cIp

Complete assignment of all unambiguous ^{1}H and ^{15}N chemical shifts of the cIp peptide was performed based solely upon homonuclear 2D ^{1}H-DIPSI, ^{1}H-ROESY and natural abundance heteronuclear {^{1}H, ^{15}N}-HSQC NMR spectroscopy experiments (see *Table III-I*), and are detailed in Appendix D-1. A total lack of long-distance NOE restraints was observed in the ROESY spectrum, which can be viewed in Figure III-1 as a DIPSI/ROESY overlay of the 'fingerprint' region. There were no d_{NN}, $d_{\alpha N(i, i+3)}$, $d_{\alpha\beta(i, i+3)}$ or any other common secondary structural NOEs (β-sheet or α-helix) observed for this peptide in solution, yielding the conclusion that it is completely unstructured in solution. This is not surprising when considering the large amount of basic and relative lack of hydrophobic residues. The poor dispersion of the amide chemical shifts in the {^{1}H, ^{15}N}-HSQC are in agreement with a conclusion of a random coil conformation. A reduction from 30°C to 5°C made no substantial changes to the NMR spectra and no low-temperature induced structural formation was observed. The peptide chemical shifts did not change over the pH range 4.0 to 7.0.

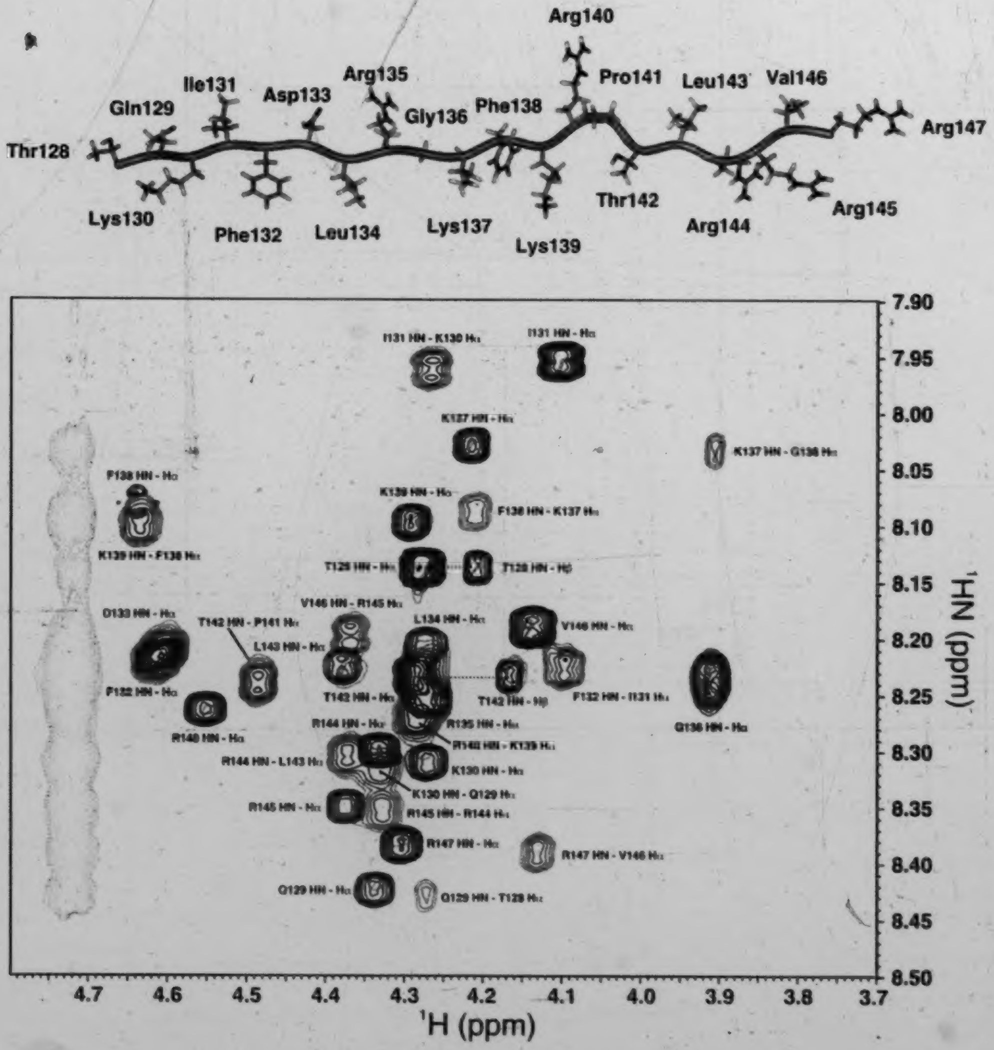


Figure III-1. **Unbound cIp in solution.** 2D ^{1}H-^{1}H-homonuclear superimposed, assigned NMR spectra of the HN-Hα region of DISPI (red) and ROESY (blue) reveals an unstructured conformation of cIp in solution. The random coil structure of the cIp region is shown above, with numbering and labeling of individual amino acids.

Table III-I

NMR spectra acquired and experimental conditions used for clp peptide

Experiment Name	Nuclei[a]	^{1}H freq	nt[b]	*x*-pts[c]	*y*-pts	*x*-sw	*y*-sw	Mix[e]
2D - {^{1}H,^{15}N}-HSQC	^{1}H,^{15}N	500	64	1024	512	6000	1800	
2D - {^{1}H, ^{1}H }-DIPSI	^{1}H, ^{1}H	600	32	4096	256	5000	5000	30.0
2D - {^{1}H, ^{1}H }-ROESY	^{1}H, ^{1}H	600	32	4096	256	5000	5000	80.0
2D - {^{13}C/^{15}N}-filtered DIPSI	^{1}H(^{14}N/^{12}C), ^{1}H(^{14}N/^{12}C)	500	32	4096	256	6000	6000	43.2

[a] The nucleus acquired in each dimension.

[b] The number of transients acquired for each free induction decay.

[c] *x,y*-points and sw are the number of complex points and sweep width in each respective dimension (*x* is the directly detected dimension).

[e] Mixing times are given in milliseconds. In the case of DIPSI/ROESY experiments, this is the spin-locking time.

III-B. Induced chemical shift changes to ^{1}H-clp upon binding to the C-lobe

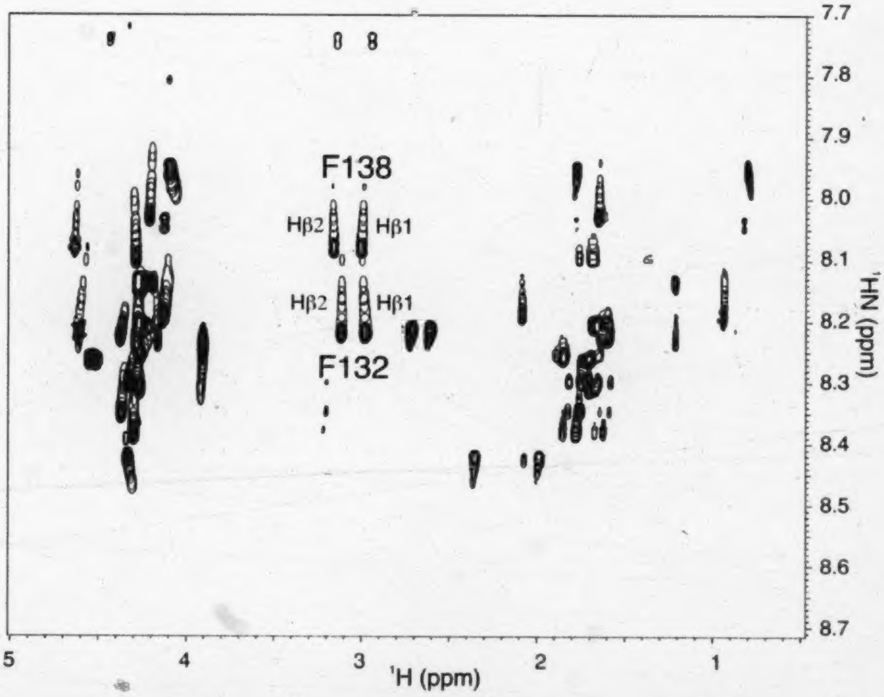


Figure III-2. **Titration of ^{1}H-clp with C-lobe**. The progressive movement of peaks can be viewed in this 2D-*NC*-filtered spectrum. Chemical shift changes of the Hβ's of residues Phe132 and Phe138 are indicated.

To determine the chemical shifts of clp in the presence of the C-lobe, ^{1}H-chemical shifts of clp were monitored during the addition of ^{13}C/^{15}N-cCTnC in a 2D-*NC*-filtered

NMR experiment (see *Table III-1)*. The *NC*-filtered experiment was useful to observe only the protons bound to unlabeled protein (i.e. ^{1}H-cIp), whereas proton resonances associated within the labeled C-lobe were not observed. As can be seen in Figure III-2, there were significant shifts in the spectrum, with diminishing peak intensity as increasing amounts of C-lobe was added to the NMR tube. The resonances of the Hβ protons of Phe132 and Phe138 were particularly useful in identification of inter-domain hydrophobic NOEs in subsequent NOE-based structural experiments for this complex. Initially at near-identical chemical shifts, there is a noticeable decrease in the splitting between the Hβ resonances for Phe132, and an increase in the splitting for the Hβ resonances for Phe138 upon binding of the C-lobe. The differences in these respective chemical shifts for the bound peptide were used as starting points for hydrophobic NOE assignments between the two respective domains in subsequent heteronuclear NMR experiments, as these chemical shifts were in a non-crowded area of the ^{1}H dimension and were unique, sharp resonances when compared to the rest of the complex.

Upon the identification of the ^{1}H-chemical shifts of the bound cIp region to the C-lobe, an *NC*-chirp experiment was run on this sample in an attempt to ascertain inter-domain NOE's between the two bound proteins. Unfortunately, there was very little signal to noise in this experiment and no usable data was obtained for this sample. Hence for the generation of unambiguous inter-domain NOEs for structural determination, the presence of inhibitory peptide that contained ^{15}N and/or ^{13}C was a necessity. The ability to ascertain intra-peptide NOEs was an added bonus in the assignment of peptide restraints for structural calculations.

III-C. Titration of ^{15}N-cIp with cCTnC•2Ca^{2+}

Monitoring of the titration of ^{15}N-cIp with ^{1}H-cCTnC revealed chemical shift perturbation, with the largest degree of perturbation in the central region of cIp peptide (Phe132 to Lys139). Residues undergoing significant chemical shift perturbation are labeled on Figure III-3. Two species of homoserine (hSer) are observed, showing resonances of hSer and hSer-lactone species present on the C-termini of the recombinant peptide. However, neither of the peaks arising from the hSer and/or hSer-lactone undergo chemical shift perturbation, indicating that the C-termini residues of cIp are not involved in binding of the complex.

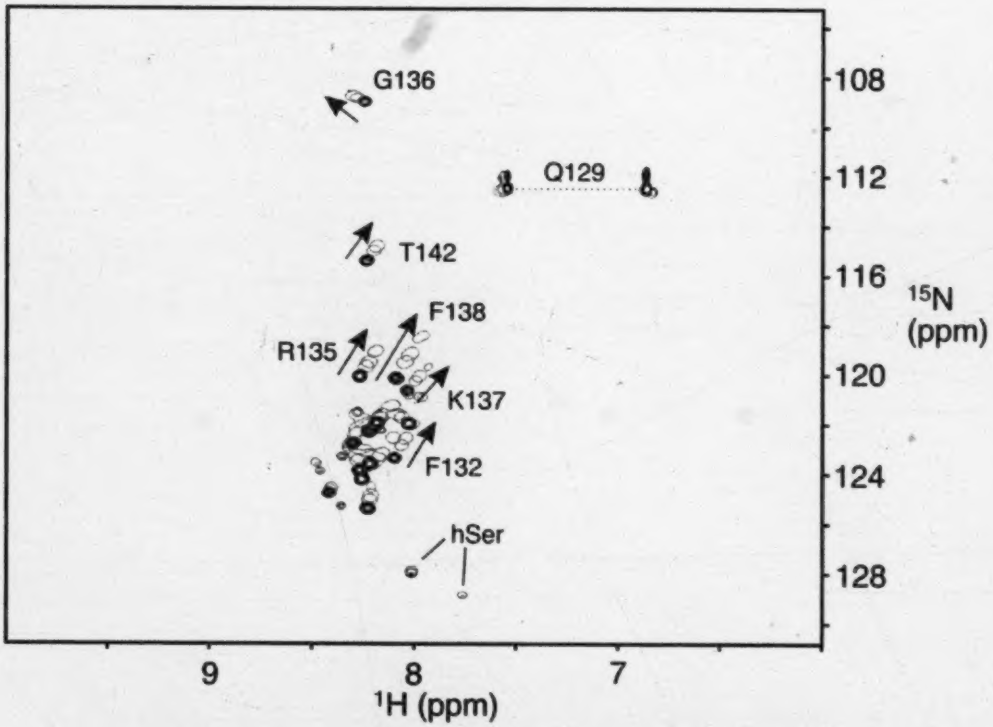


Figure III-3. **Titration of ^{15}N- cIp with ^{1}H-cCTnC**. Two-dimensional {^{1}H, ^{15}N}-HSQC NMR spectra arising from the backbone NH and side chain NH_2 groups of ^{15}N-cIp (●) at various concentrations of ^{1}H-cCTnC are superimposed (○), showing the progressive shifts of peaks. The direction of chemical shift changes for some residues undergoing large perturbations are indicated by an arrow.

The 2D {^{1}H, ^{15}N}-HSQC NMR spectrum of cIp was completely assigned and was used as a starting point to monitor the ^{1}H-cCTnC induced chemical shift changes. All of the chemical shift changes fall into the fast exchange limit on the NMR time scale. Linear movement of cross-peaks indicates that only two species exist and the binding occurs with a 1:1 stoichiometry in the interaction between cCTnC•2Ca^{2+} and the respective cIp peptide. Each cross peak position corresponds to the weighted average of the bound and free chemical shifts of cCTnC•2Ca^{2+}. The relative lack of amide chemical shift dispersion upon binding of the C-lobe, coupled with the induced chemical shift perturbations, indicate that the bound structure of the inhibitory peptide may contain unstructured regions in the C-terminal tail from resonances Arg140 to Arg147. An {^{1}H, ^{15}N}-HSQC overlay of the unbound synthetic peptide (^{15}N-natural abundance) with the recombinant ^{15}N-cIp revealed identical spectra (with the exception of the N-terminal amide group resonances of the synthetic peptide). It can therefore be noted that the presence of the hSer on the C-terminal end of the peptide does not affect the conformation in solution, as well does not affect a change from the 'capped/protected' synthetic peptide.

III-D. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp

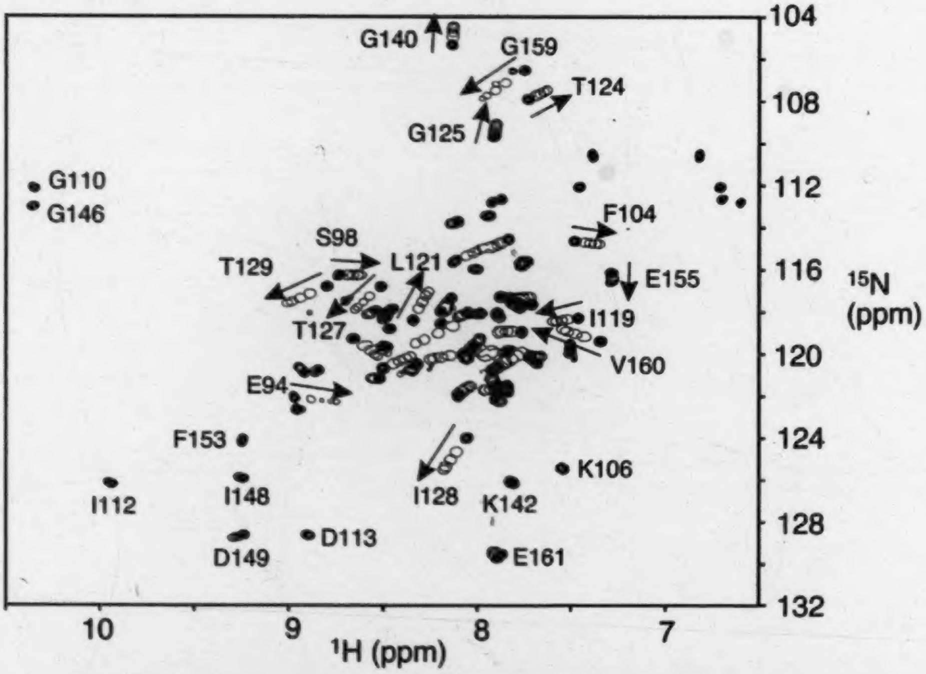


Figure III-4. **Titration of ^{15}N-cCTnC with ^{1}H-cIp**. Two-dimensional {^{1}H, ^{15}N}-HSQC NMR spectra arising from the backbone NH and side chain NH_2 groups of ^{15}N-cCTnC•2Ca^{2+} (●) at various concentrations of cIp are superimposed (○), showing the progressive shifts of peaks. The direction of chemical shift changes for some residues undergoing large perturbations are indicated by an arrow.

To obtain the bound chemical shifts of the C-lobe in the presence of synthetic cIp peptide, a titration of ^{15}N-cCTnC•2Ca^{2+} with ^{1}H-cIp was performed and monitored by 2D {^{1}H, ^{15}N}-HSQC NMR spectroscopy, with chemical shift changes of the C-lobe shown in Figure III-4. The 2D {^{1}H, ^{15}N}-HSQC NMR spectrum of unbound cCTnC•2Ca^{2+} has been previously assigned (*5*) and was used as a starting point to monitor the cIp peptides induced chemical shift changes. The chemical shift changes for all resonances fall into the fast exchange limit on the NMR time scale. Linear movement of cross-peaks indicates that only two species exist and binding occurs in a 1:1 stoichiometry. Each cross peak position corresponds to the weighted average of the bound and free chemical shifts of cCTnC•2Ca^{2+}. Superimposition of the {^{1}H, ^{15}N}-HSQC NMR spectra of ^{15}N-cCTnC•2Ca^{2+} saturated with either the synthetic and recombinant cIp peptides revealed that the spectra were identical, indicating that both recombinant and synthetic cIp peptides used in this study perturb the C-lobe in an identical manner. Therefore the presence of the hSer on the C-terminal tail of the recombinant peptide does not affect binding the C-lobe.

Figure III-4 reveals the differing degrees of induced chemical shift changes for different residues in cCTnC•2Ca^{2+} upon cIp binding. For example, residues such as Gln122 and Val160 undergo large chemical shift perturbations, indicating potential contacts with cIp, whereas residues such as Ile112 and Asp113 show relatively small cIp-induced chemical shifts. Thus, it is useful to correlate peptide induced chemical shift changes to residues along the cCTnC•2Ca^{2+} protein sequence. Figure III-5 plots the absolute chemical shift changes in both ^{1}H and ^{15}N dimensions against the protein sequence. This plot shows that many residues experience significant peptide induced chemical shift perturbations ($\Delta\delta_{ppm} > 0.5$ ppm). The mapping of these induced chemical shifts are displayed in a surface plot representation of the structure of the C-lobe in Figure III-6.

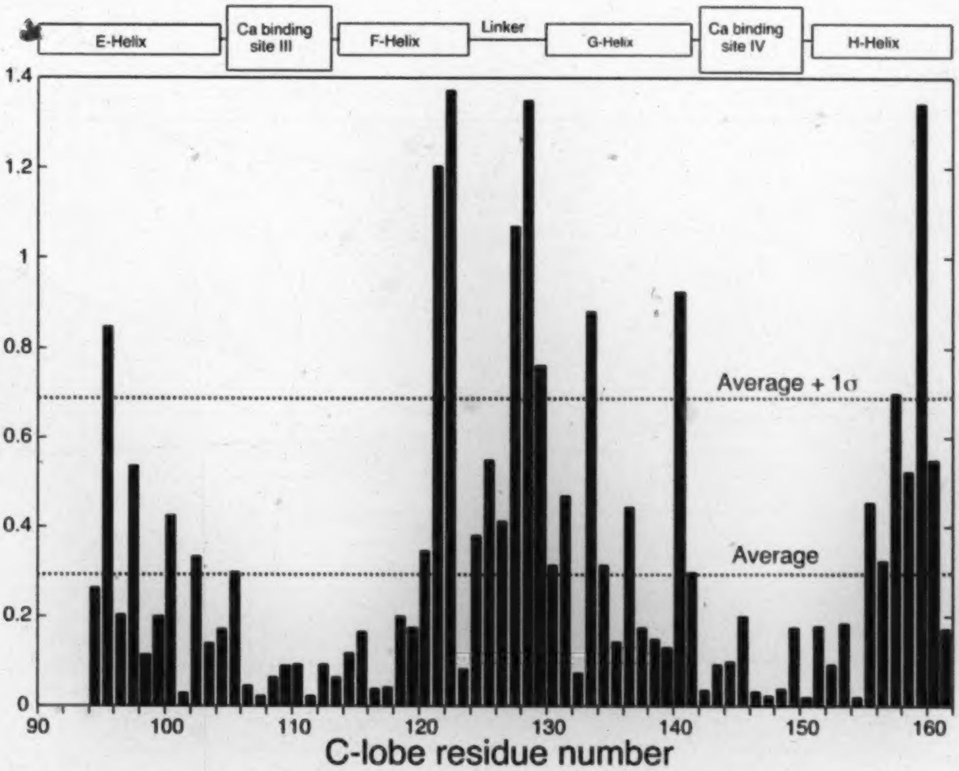


Figure III-5. **Comparison of induced chemical shift changes of cCTnC•Ca^{2+} upon the addition of cIp peptide.** The chemical shift values are reported as the square root of the sum of the squares of chemical shift changes for both the ^1H and ^{15}N dimensions in the HSQC spectra. [$\Delta\delta(\text{ppm}) = (\Delta\delta N^2 + \Delta\delta H^2)^{1/2}$]

Resonances that underwent significant backbone amide ^1H and/or ^{15}N chemical shift changes during titration were followed to monitor peptide binding to protein. For cIp titration, the peptide induced chemical shift changes of individual amide resonance as a function of the $[\text{cIp}]_{\text{total}}/[\text{ cCTnC}\bullet 2\text{Ca}^{2+}]_{\text{total}}$ ratio was plotted, and the average curve is shown in Figure III-7. The normalized average chemical shift data as a function of $[\text{cIp}]_{\text{total}}/[\text{ cCTnC}\bullet 2\text{Ca}^{2+}]_{\text{total}}$ were fit to equation [1] using the software program XCRVFIT (*50*), which yielded a dissociation constant (K_D) of 31 ± 11 *u*M.

$$\text{cCTnC}\bullet 2\text{Ca}^{2+} + \text{cIp} \leftrightarrow \text{cCTnC}\bullet 2\text{Ca}^{2+}\bullet\text{cIp} \quad [1]$$

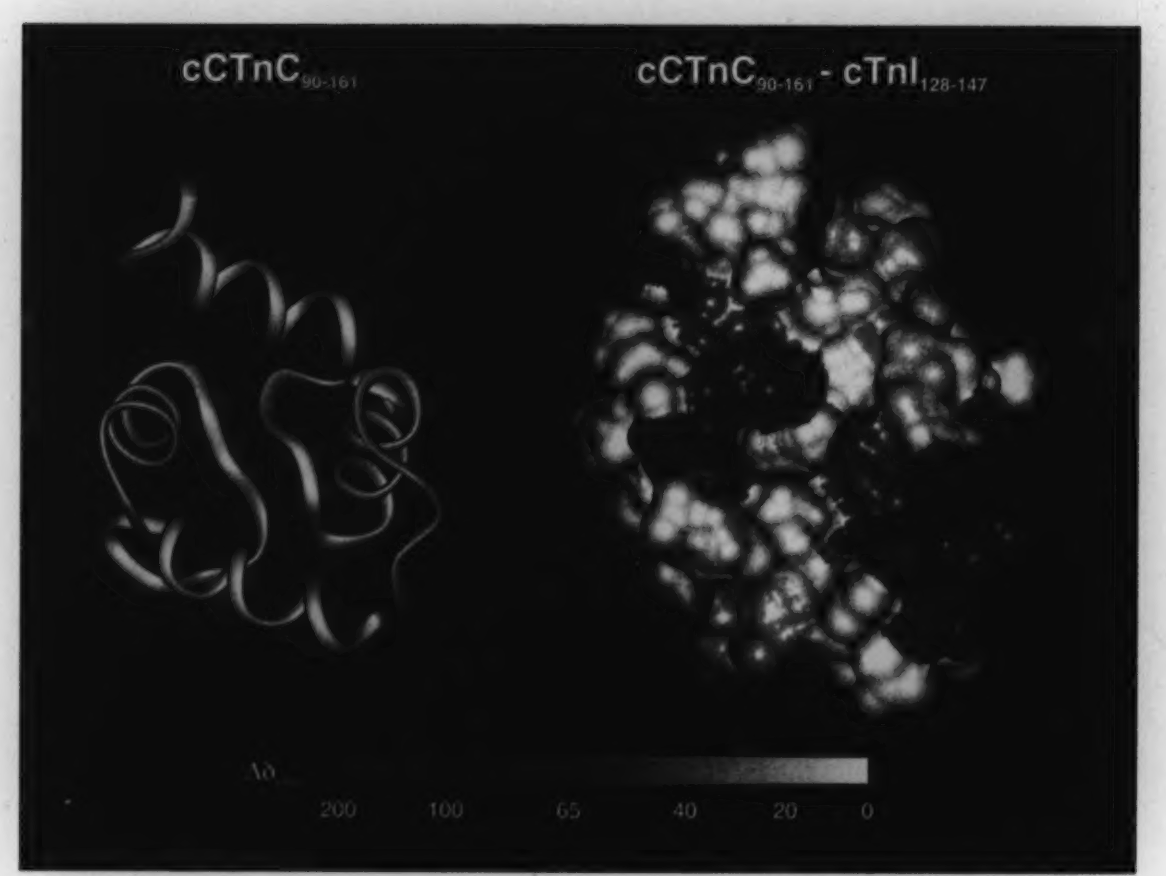

cCTnC90-161
cCTnC90-161 - cTnI128-147
Δδ
200
100
65
40
20
0

Figure III-6. **Chemical shift map of the C-lobe during cIp binding.** A ribbon diagram of the C-lobe in the unbound form is shown to the left to establish an orientation for the surface rendering to the right. Resonances that undergo large chemical shift perturbations (in Hz) are colored in red and minimal perturbations are colored in white.

For the generation of a binding constant, only chemical shift changes for resonances Ile128, Thr129, Gly159 and Val160 were used. The induced chemical shifts for each resonance for each titration point were measured against their respective maximal induced chemical shift (saturation) and were each normalized to 1.0 individually. The normalized values for the four resonances were then averaged prior to obtaining a dissociation value for the titration. It was found that the averaging of the four resonances produced a binding constant with error that was virtually identical to that if 10, 20 or 30 different resonances were averaged. The four resonances were chosen as they were in spectral areas where no chemical shift overlap occurred and they exhibited shifts in the both the ^{15}N and ^{1}HN dimensions, so as not to bias the data. Therefore only 4 resonances were used in the determination of a binding constant for this titration. This became a large time saving method in Chapter IV of this thesis where a glycine-scanning mutagenesis of the cIp peptide is performed to obtain thermodynamic binding data for electrostatic investigations.

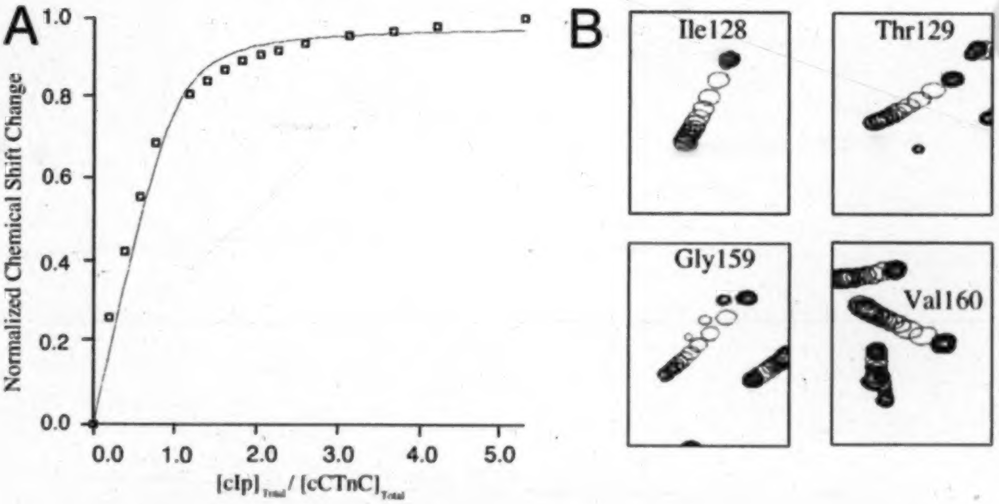


Figure III-7. **Dissociation curve of cIp binding to the C-lobe**. The normalized chemical shift change is an average for resonances Ile128, Thr129, Gly159 and Val160 shown in Figure III-4, all of which underwent perturbations greater than one standard deviation of the mean.

III-E. Overall Structure of the cCTnC•2Ca^{2+}•cIp complex

The NMR experiments performed to obtain structural data for the TnC-TnI complex are summarized in *Table III-II.* A sample of ^{13}C/^{15}N-cCTnC in complex with ^{1}H-cIp gave a 2D {^{1}H, ^{15}N}-HSQC NMR spectrum that was highly resolved allowing the assignment of chemical shifts of the cCTnC backbone and side chain nuclei. Previously, our laboratory has obtained data for protein-peptide/ligand complexes using ^{13}C/^{15}N-labeled protein in complex with unlabeled peptides/ligands (*51-53*), using two-dimensional ^{13}C/^{15}N-filtered DIPSI and NOESY experiments (*54*) to obtain chemical shift data of the bound peptide/ligand in the complex (see Figure III-2). Production of ^{13}C/^{15}N-cIp peptide was essential for this study as information obtained from standard filtered experiments produced very limited information, suggesting few close hydrophobic contacts, which would contribute to NOE cross peaks, occur between the protein and the peptide in the binary complex. Edited experiments of ^{13}C/^{15}N-cIp complexed with unlabeled cCTnC allowed for complete chemical shift assignment of cIp atoms in the bound binary complex. Three samples were used in this study to obtain NMR structural restraint data for the binary complex; ^{13}C/^{15}N-cCTnC•2Ca^{2+}•cIp, cCTnC•2Ca^{2+}•^{13}C/^{15}N-cIp, and ^{13}C/^{15}N-cCTnC•2Ca^{2+}•^{13}C/^{15}N-cIp. Initially, inter-domain NOEs were assigned starting with the known, and unique chemical shifts for the Hβ resonances for the two Phe side-chain (Phe132 and Phe138) of cIp. These were used as a starting point and as a reference point for inter-domain contacts. Distance restraints for cCTnC and cIp in the complex were obtained by analyzing three-dimensional ^{15}N-, ^{13}C-, and/or ^{13}C/^{15}N-edited NOESY experiments. Dihedral angle restraints for cCTnC and cIp in the complex were obtained from three-dimensional HNHA experiments and ^{15}N-edited NOESY experiments. Sample ^{13}C/^{15}N-cCTnC•2Ca^{2+}•^{13}C/^{15}N-cIp was used for assignment of symmetrical peaks in the ^{13}C-edited NOESY experiment for unambiguous protein-peptide side chain NOE cross peak assignments.

A total of 1313 experimental distance restraints were obtained for the complex and used to calculate the high resolution solution structure of the complex of cCTnC•2Ca^{2+}•cIp: 1046 intramolecular NOE distance restraints for cCTnC (~15

Table III-II

NMR spectra acquired and experimental conditions used to obtain assignments and NOE restraints

A. Sample $^{13}C/^{15}N$-cCTnC complexed with cIp

Experiment Name	Nuclei[a]	^1H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D - {$^1H,^{13}C$}-HSQC	$^1H,^{13}C$	500	16	876	128		6000	10056			*(55)*
2D - {$^1H,^{15}N$}-HSQC	$^1H,^{15}N$	500	32	1024	128		6000	1800			*(56)*
3D - ^{15}N-edited DIPSI-HSQC	$^1H,^1H,^{15}N$	600	16	900	114	30	7500	6300	1600	80.2	*(57)*
3D - ^{15}N-edited NOESY-HSQC	$^1H,^1H,^{15}N$	500	16	640	128	32	6000	5250	3000	80	*(57)*
HNHA	$^1H,^1H,^{15}N$	600	16	960	48	48	7500	4807	1700		*(58)*
CBCA(CO)NNH	$^1H,^{13}C,^{15}N$	500	20	768	50	32	6000	7794	1450		*(59)*
HNCACB	$^1H,^{13}C,^{15}N$	500	16	888	70	32	6000	7794	1450		*(59)*
HCCH-TOCSY	$^1H,^1H,^{13}C$	500	16	768	128	32	6000	3200	3000	19.6	*(60)*
3D - $^{13}C/^{15}N$-edited NOESY	$^1H,^1H,^{13}C/^{15}N$	500	16	640	128	32	6000	5250	3000	80	*(61)*

B. Sample $^{13}C/^{15}N$-cIp complexed with cCTnC

Experiment Name	Nuclei[a]	^1H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D - {$^1H,^{13}C$}-HSQC	$^1H,^{13}C$	500	16	876	128		6000	10056			*(55)*
2D - {$^1H,^{15}N$}-HSQC	$^1H,^{15}N$	500	32	1024	128		6000	1800			*(56)*
3D - ^{15}N-edited DIPSI-HSQC	$^1H,^1H,^{15}N$	600	24	922	96	32	7200	5681	1800	80.2	*(57)*
3D - ^{15}N-edited NOESY-HSQC	$^1H,^1H,^{15}N$	600	24	922	96	32	7200	5681	1800	80	*(57)*
HNHA	$^1H,^1H,^{15}N$	600	16	922	96	32	7200	5681	1800		*(58)*
HNCACB	$^1H,^{13}C,^{15}N$	500	32	1024	50	32	6000	7794	1650		*(59)*
HCCH-TOCSY	$^1H,^1H,^{13}C$	500	16	768	128	32	6000	3000	3000	19.6	*(60)*
3D - $^{13}C/^{15}N$-edited NOESY	$^1H,^1H,^{13}C/^{15}N$	500	16	600	128	32	6000	6000	3000	75	*(61)*
3D - ^{13}C-edited NOESY	$^1H,^1H,^{13}C$	800	16	1408	100	32	12000	9118	6000	80	*(61)*

C. Sample $^{13}C/^{15}N$-cCTnC complexed with $^{13}C/^{15}N$-cIp

Experiment Name	Nuclei[a]	^1H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
3D - ^{13}C-edited NOESY	$^1H,^1H,^{13}C$	800	16	1344	128	32	12001	10000	8000	80	*(61)*

[a] The nucleus acquired in each dimension (i.e. $^1H,^{15}N$ indicates ^1H in x-dimension and ^{15}N in y-dimension).

[b] The frequency of ^1H in MHz.

[c] the number of transients acquired for each FID

[d] x,y,z-pts and sw is the number of complex points and sweep width in each respective dimension (x is the directly detected dimension)

[e] mixing times are given in milliseconds. In the case of DISPI experiments, this is the spin-locking time.

restraints per residue), 267 intramolecular NOE distance restraints for cIp (~13 restraints per residue), 23 intermolecular NOE distance restraint between cCTnC and cIp, 29 dihedral restraints for cCTnC and 12 cCTnC distance restraints to Ca^{2+}. Figure III-8 depicts the ensemble of the 30 lowest energy structures of the complex with superimposition of the backbone heavy atoms of $cCTnC_{95-155}$. The overall conformational energies and structural statistics for the ensemble are provided in *Table III-III*. A ribbon diagram of the lowest energy structure of the ensemble is provided in Figure III-8, and a surface map depicting the perturbation of amide resonances of cCTnC upon cIp binding is additionally provided. The structures of cCTnC in the ensemble are of higher quality than those of cIp, as a result of more restraints per residue for the protein than for the peptide. The majority of assigned cIp intramolecular NOEs in the complex are of intraresidue origin, with no long range ($i - j \geq 5$) NOEs observed for cIp.

Complete chemical shift assignments for both the C-lobe and the inhibitory peptide in the binary complex are listed in Appendix D-2.

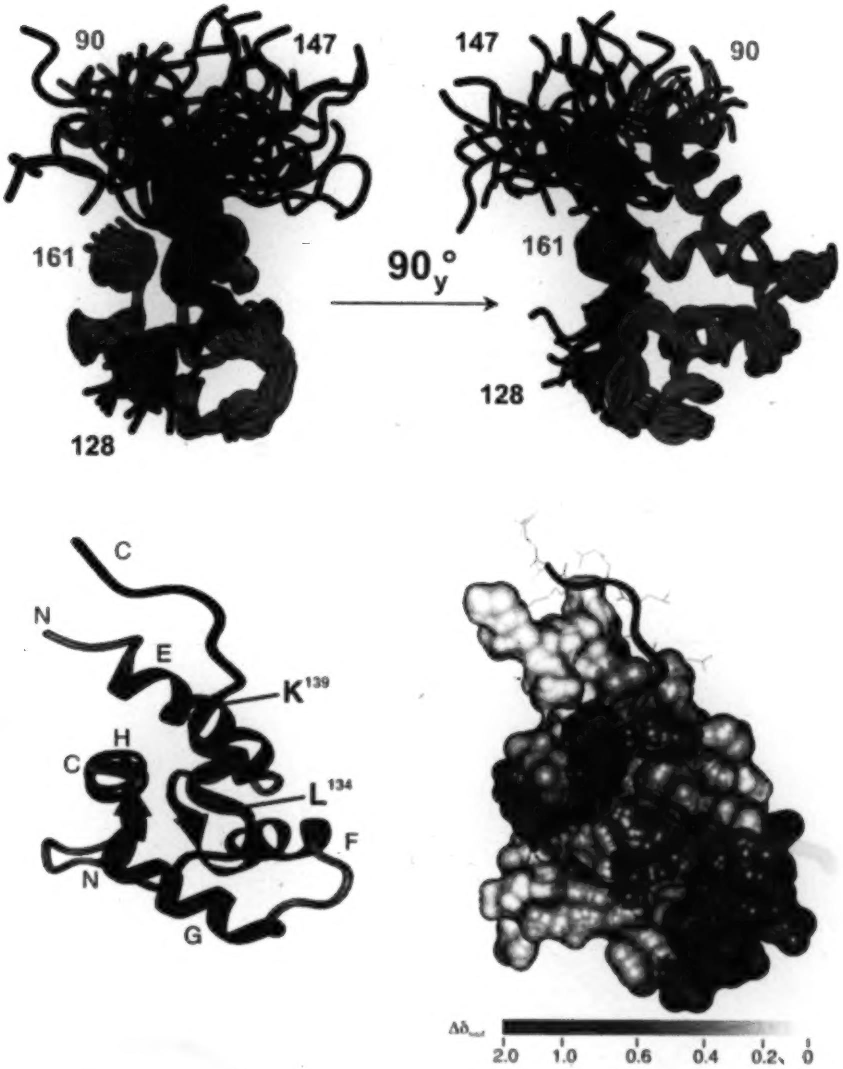


Figure III-8. **The structure of the bound inhibitory region with the C-lobe**. An ensemble plot of the 30 lowest energy structures of cCTnC•2Ca^{2+} (green) in complex with cIp (blue), via super-imposition of the C-lobe backbone [top] (1OZS.pdb). Residues Ile131-Lys139 are well ordered, while residues Arg140-Arg147 are flexible. Ribbon diagram of the binary complex cCTnC•2Ca^{2+}•cIp [bottom-left]. The helical region of cIp encompasses residues Leu134-Lys139 (red) completing 1.5 turns of a helix is colored in red. N- and C- termini of each protein in the binary complex are as indicated, as well all helices of cCTnC are defined. Chemical shift surface map of cCTnC•2Ca^{2+}•cIp as based upon the titration performed in Figure III-4 [bottom-right]. cIp peptide backbone is colored as blue, with all heavy side chain atoms colored as green.

Table III-III

Structural statistics of the family of the 30 structures calculated

NOE restraints	**cCTnC**	**cTnI$_{128-147}$**	**Complex**
Total	1046	267	1313
Intraresidue	444	178	622
Sequential ($\|i - j\| = 1$)	352	72	424
Medium Range ($2 \leq \|i - j\| \leq 4$)	143	17	160
Long Range ($\|i - j\| \geq 5$)	107	0	107
Artificial restraints to Ca^{2+}	12	0	12
Dihedral restraints			
ϕ	29	0	29
Intermolecular NOEs between cCTnC and cIp			23

r.m.s.d. from the average structure (Å)	Backbone atoms	Heavy atoms
cCTnC[a]		
Well-defined residues	0.58 ± 0.12	0.62 ± 0.13
Helices	0.33 ± 0.10	0.35 ± 0.11
cTnI$_{128-147}$[b]		
Helical region	0.45 ± 0.12	0.78 ± 0.21

Energies (kcal/mol)	Complex	**Restraint Violations**	Complex
E_{total}	88.8 ± 5.4	Distance > 0.2 Å[c]	0
E_{NOE}[d]	3.6 ± 0.8	Dihedral > 1°	0
		ϕ, ψ in core or allowed regions[d]	96.6%

[a] 55 residues have well defined backbone atoms. This includes residues 94-122, 131-158. The r.m.s.d. for helices are calculated from the average of the r.m.s.d. of individual helices.

[b] Superimposition of helical region Leu134-Lys139

[c] There were no distance restraints over 0.2 Å.

[d] The final K_{NOE} used was 50 kcalmol^{-1}

[e] As determined using PROCHECK (_47_) – refer to Appendix D-3.

III-F. Structure of cCTnC in the complex

The overall fold of cCTnC in the complex resembles other Ca^{2+}-bound domains in the EF-hand family, such as the C-lobe of skeletal TnC (_1_). The secondary structure elements of cCTnC in the complex are identical to unbound cCTnC•2Ca^{2+} (_5_), displaying four well-defined helices (helices E, F, G, and H) and two well-defined anti-parallel β-sheets. This helix-loop-helix structural motif is common in many Ca^{2+} binding proteins

(Ca^{2+} atoms are not shown in Figure III-8). For the 30 calculated lowest energy structures, the four helices superimpose with individual backbone r.m.s.d. values of 0.34 ± 0.10 Å for helix E (residues 92-105), 0.30 ± 0.07 Å for helix F (residues 114-123), 0.30 ± 0.13 Å for helix G (residues 130-140) and 0.27 ± 0.10 Å for helix H (residues 150-160). The two anti-parallel β-sheets (residues 111-113, 147-149) superimpose with backbone r.m.s.d value of 0.09 ± 0.04 Å. Refer to Appendix D-3 for a complete procheck profile of the ensemble. The N-terminal residues (89-91) are not as well defined, yet unexpectedly the C-terminal residues (159-161) are well ordered when compared to the unbound structure (*5*). This ordering of the C-terminal residues is due to NOE contacts with clp in the complex, with contacts between clp-Phe138 and cCTnC-Val160 minimizing the flexibility of the C-terminal residues, yielding a more ordered H-helix near residues 159-161 (refer to Figure III-9 for an example of NOE connectivities).

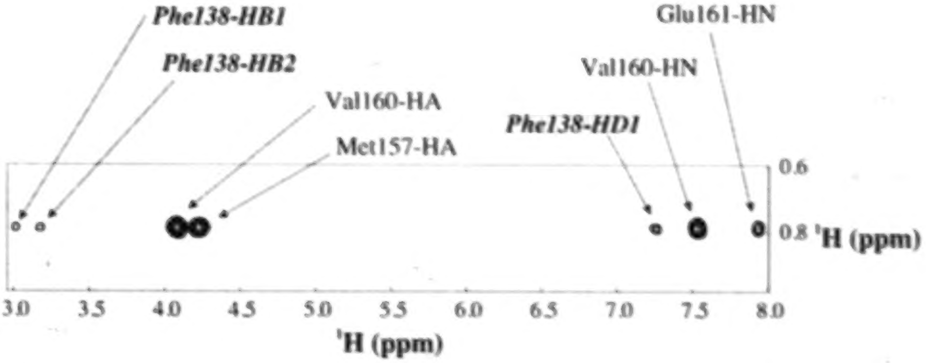


Figure III-9. **^{13}C-NOESY strip-plot of Val160 connectivities.** NOE plot of Val160-HG21 (^{13}C=20.785ppm) contacts with clp-Phe138 from ^{13}C-edited NOESY NMR experiment. Inter-domain contacts from cCTnC-Val160 to clp-Phe138 of are indicated ***in italics***.

EF hand structures can be described as 'opened' or 'closed', based upon the interhelical angles that the E-F and G-H helices make with respect to one another. A comparison of interhelical angles of cCTnC in the binary complex with previously solved EF hand structures is necessary to judge the changes in the C-lobe backbone orientation upon clp binding, and is presented in *Table III-IV*. The binding of cSp to calcium saturated cNTnC reveals a conformational opening of over 20°, yet binding of clp to calcium saturated cCTnC reveals only a slight conformational opening. Comparison with the x-ray crystal structure of $sTnC{\bullet}2Ca^{2+}{\bullet}sTnI_{1-47}$ reveals identical interhelical angles of the G-H helices, and a slight closing of the E-F helix. Thus it is concluded that unlike the N-lobe of cTnC, the C-lobe is unable to undergo large conformational opening/closing upon target peptide binding.

Table III-IV

Interhelical angles of various EF hands

Calcium binding protein	Interhelical angles (°)[a]		Accession code (PDB)	Ref.
	E-F	G-H		
cCTnC • $2Ca^{2+}$ (NMR)	115 ± 4	121 ± 4	1AJ4	(*5*)
cCTnC • $2Ca^{2+}$ • $cTnI_{128-147}$ (NMR)	108 ± 4	108 ± 4	1OZS	(*62*)
cNTnC • apo (NMR)	136 ± 3[b]	129 ± 5[b]	1SPY	(*63*)
cNTnC • Ca^{2+} (NMR)	132 ± 3[b]	117 ± 3[b]	1AP4	(*63*)
cNTnC • Ca^{2+} • $cTnI_{147-163}$ (NMR)	102 ± 5[b]	96 ± 5[b]	1MXL	(*52*)
sTnC • $4Ca^{2+}$ (NMR)	105 ± 6	111 ± 7	1TNW	(*64*)
sTnC • $2Ca^{2+}$ • $sTnI_{1-47}$ (x-ray)	99	109	1FI5	(*18*)
cCTnC • $2Ca^{2+}$ • EMD57033 (NMR)	92 ± 4	113 ± 7	1IHO	(*53*)

[a] A large angle defines a closed conformation, whereas a small angle defines an open conformation.

[b] For interhelical angles of the N-terminal domain of troponin C, helices E-F refer to helices A-B and G-H to C-D respectively. The axis for an a-helix is defined by two points, taken as the average coordinates of the first and last 11 backbone atoms of the a-helix.

III-G. Structure of cIp in the complex

The structure of cIp in the binary complex displays a region of helical content in the central region of the peptide, flanked by a non-helical structured region on the N-termini and an unstructured random coil region on the C-termini. The overall fold of cIp in complex with cCTnC is novel, possessing no structural similarities to known Ca^{2+}-binding protein/peptide interactions. cIp runs anti-parallel to cCTnC in the complex, with residues Ile131-Asp133 placed in the hydrophobic cleft of cCTnC. Only Phe132 makes definitive hydrophobic NOEs with the C-lobe. Contacts to Ile131 were ambiguous and thus not included in the final calculation of structures. The observation that Ile131 is within the hydrophobic core is merely an artifact of the cns anneal protocol. Residues Ile131-Asp133 possess no definitive helical secondary structure, yet superimposition of the 30 lowest energy (Figure III-8) structures reveals a structured region. This structured region makes a near 90° turn along the hydrophobic cleft of cCTnC to begin the helical region beginning at residue Leu134. Multiple hydrophobic NOE contacts were observed Phe132 making contacts with cCTnC-Leu121/Ala123/Ile128. Following the N-terminal residues, Leu134-Lys139 adopt a helical secondary structure, with NOE contacts of cIp-Phe138 making hydrophobic contacts with cCTnC-Val160/Leu100 (E and H helices, see Figure III-9 for an example), with E helix contacts closely matching those predicted by Cachia and coworkers for the inhibitory region (*65*). Checking of the helical region spanning residues Leu134-Lys139 by the program Procheck (*47*) for the 30 calculated lowest energy structures reveals various violations of the backbone dihedral angles in comparison with accepted Ramachandran values, indicating a non-ideal helix (see Appendix D-3). The structure of the inhibitory region could change in the troponin complex containing full length TnC, TnI and TnT, wherein additional stabilizing protein-protein interactions may be present. For the 30 calculated lowest energy structures, the central helix of cIp superimposes with individual backbone r.m.s.d. values of 0.45 ± 0.12 Å for residues Leu134-Lys139. Residues Arg140-Arg147 comprising the C-terminal region of cIp display no evidence of a well-ordered secondary structure. Superimposition of the 30 calculated lowest energy structures of the binary complex (see Figure III-8) reveals this region as very labile in solution. No long-range peptide-peptide NOE contacts were observed for this region, with only *i*, *i*+1 contacts observed between residues, as well no protein-peptide NOE contacts were observed for this region. Arg147 begins the segment of the 'switch peptide' (cSp, residues 147-163), which has been shown to start a helical secondary structure when complexed with cNTnC•Ca^{2+} (*51, 52, 66*). Arg147 makes multiple contacts with cNTnC, and thus is not expected to make interactions with cCTnC in the complex. Since the hSer on the C-terminal tail of the recombinant peptide is C-terminal to Arg147, it is not at all surprising that earlier data reported in Figure III-4 showed identical C-lobe chemical shift values for the synthetic (no hSer) vs. the recombinant peptide.

Figure III-8C displays the same chemical shift surface map of the C-lobe as was displayed in Figure III-6, however cIp is present in this representation. No chemical shifts were measured for Lys90-Ser93 due to exchange with water, and are thus reported as no change in chemical shift and are colored white. The regions of greatest chemical shift change are viewed in red, with largest chemical shift perturbations of cCTnC upon the binding surface area of cIp, within the linker region and the E and H helices. A large chemical shift was observed for the linker region, as Phe132 from the inhibitory region is within close proximity to the linker region, with aromatic ring effects predicted to be responsible for the large shifts observed.

Chemical shift values for $^{13}C_\alpha$ and $^{1}H_\alpha$ nuclei are important indicators for predicting secondary structure (*67-69*). NMR studies have shown that upon the initiation of helical secondary structure formation, there is a downfield shift observed for $^{13}C_\alpha$ and an upfield shift observed for $^{1}H_\alpha$ (thus $\Delta\delta(\text{ppm}) = \delta_{\text{bound}} - \delta_{\text{free}}$). Incorporation of ^{13}C label in cIp allows for measurement of both $^{13}C_\alpha$ and $^{1}H_\alpha$ chemical shifts of free and bound peptide. cIp undergoes $^{13}C_\alpha$ chemical shift perturbation towards a more helical conformation (downfield shift; $(\delta_{\text{bound}} - \delta_{\text{free}}) > 0$) in the region of residues Leu134-Lys137, with significant shift (>1ppm) for residue Arg135. Residues Ile131-Asp133 experience small upfield changes in $^{13}C_\alpha$ chemical shift upon cCTnC binding, corresponding to the N-terminal region before the helical region of Leu134-Lys139. $^{13}C_\alpha$ chemical shift changes for residues Arg140-Arg147 indicate that this region possesses no helical characteristics, corresponding to the flexibility of this region as shown in Figure III-8. $^{1}H_\alpha$ perturbation measurements show minimal changes for all residues for cIp upon cCTnC binding, with only residues Gln129, Asp133, Arg135, Arg145 and Arg147 undergoing changes greater than 0.05ppm. For the helical region Leu134-Lys139, residues Leu134 and Arg135 undergo an upfield shift, with all others undergoing minimal downfield shifts less than 0.05ppm. Comparison with induced $^{1}H_\alpha$ chemical shift changes of the skeletal inhibitory peptide isoform (sIp) upon binding to sTnC as reported by Hernandez and coworkers yields similar results of no changes greater than 0.1 ppm for the majority of residues (*28*), with Arg135 undergoing the largest upfield chemical shift for both the cardiac and skeletal isoforms upon TnC binding.

An electrostatic surface map (GRASP rendering) of the binary complex cCTnC•$2Ca^{2+}$•cIp and its components are shown in Figure III-10. The highly basic cIp peptide has the potential for many favorable electrostatic interactions with the highly acidic cCTnC. At physiological conditions (pH 6.7), cIp (pI=11.0) will have an overall charge of 6.9, while cCTnC (pI=4.1) will possess an overall charge of −14.8 (*70*). It is expected that the highly basic C-terminal end of cIp (Arg140, Arg144, Arg145, Arg147) will make beneficial interactions with the acidic E-Helix of cCTnC (Glu94, Glu95 and Glu96), thereby stabilizing the binary complex (*65*). Therefore, it can be hypothesized that electrostatic interactions are a predominant force in the association of these two domains, as is evidenced by the relative lack of hydrophobic NOEs in the complex.

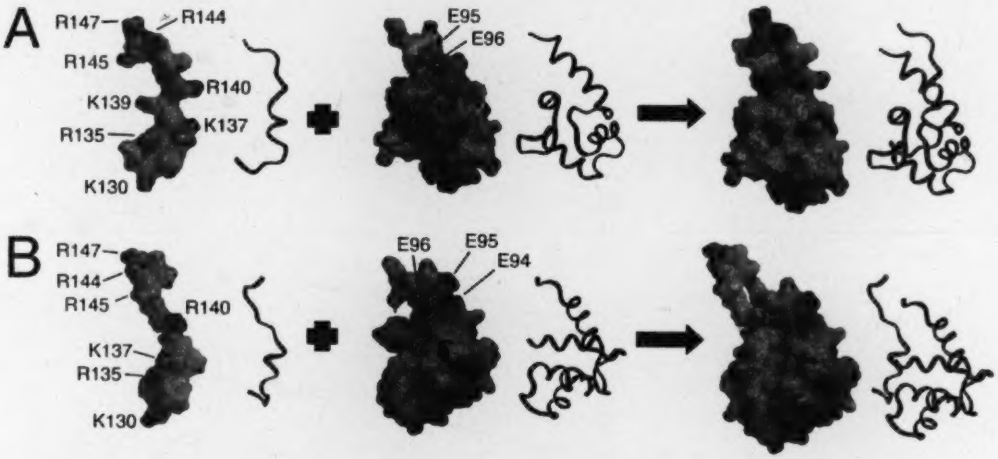


Figure III-10. **Electrostatic surface map of binary complex cCTnC•2Ca^{2+}•cIp.** (A) Ribbon diagram of domain orientation is as shown with acidic regions colored in red and basic regions colored in blue (1OZS.pdb). Selected residues within the complex are as labeled. (B) 90° rotation of complex about the y-axis. Surface maps generated using the programs GRASP and RASTER3D (*71*, *72*). Ribbon diagrams generated using the program Insight II v98 (Accelrys Inc.).

Based solely on the domain orientations shown in Figure III-10, it is easy to visual the hypothesized electrostatic interactions that might be occurring between the two domains. However, I was unable to determine the exact salt-bridging pairs. Electrostatic interactions are therefore explored in further detail in Chapter IV of this thesis, where I performed scanning glycine mutagenesis to probe the nature of the electrostatic interactions occurring in this complex.

III-H. Backbone amide ^{15}N-NMR relaxation studies of cCTnC

Backbone amide ^{15}N-NMR relaxation data for ^{15}N-cCTnC•2Ca^{2+} and ^{15}N-cCTnC•2Ca^{2+}•cIp were obtained at 500 and 600 MHz (see Figure III-11). Backbone resonances for Met90-Lys92 were not observed due to rapid amide proton exchange with water for both samples. Resonances Met131, Lys137 and Arg147 were not included in ^{15}N-cCTnC•2Ca^{2+}•cIp data, as well as resonances Ala99 and Arg147 were not included in ^{15}N-cCTnC•2Ca^{2+}, due to partial peak overlap in the {^{1}H, ^{15}N}-HSQC spectra. To avoid plot cluttering, error bars were not included as the margin of error for the majority of the residues was smaller than the size of the markers (i.e. ◆ and ○) in Figure III-11.

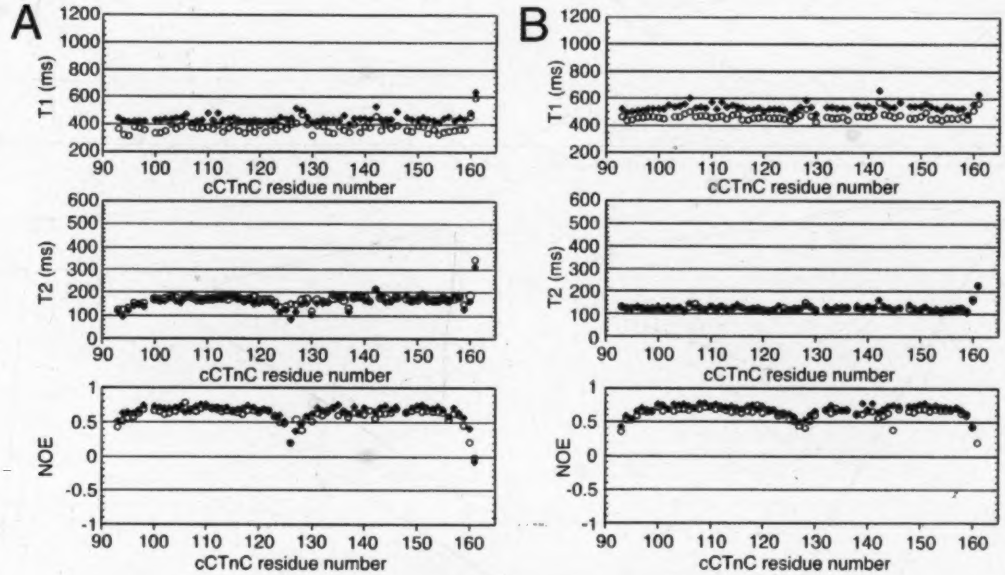


Figure III-11. **C-lobe ^{15}N-amide relaxation measurements.** ^{15}N-NMR relaxation values of ^{15}N-cCTnC for (A) unbound and (B) ^{15}N-cCTnC•2Ca^{2+}•cIp binary complex measured at 500 MHz (○) and 600 MHz (◆). T_1, T_2 values are reported in milliseconds (ms), and NOE values are the ratio of the intensity in the absence/presence of proton pre-saturation.

The measured values at 500 MHz for ^{15}N-cCTnC•2Ca^{2+} show a decrease in T_1, decrease in T_2, and increase in NOE values for the first 5 N-terminal residues on the E helix when compared to calcium saturated skeletal C-lobe isoform (sCTnC), indicating that the N-terminal residues on the E helix of cCTnC are more ordered when compared to the skeletal isoform (*32*), as well as small differences in calcium binding sites III and IV. Residues whose internal motions affect their measured relaxation values were excluded from the calculation of the averages, as determined using NOE criteria ($NOE^{500} > 0.6$ and $NOE^{600} > 0.65$). As expected, the average T_1^{600}/T_1^{500} ratio is >1 and the ratio of T_2^{600}/T_2^{500} is approximately equal to 1 for the majority of the residues studied. These ratios are within accepted theoretical calculations, taking into account effects from dipole-dipole relaxation and chemical shift anisotropy at the two magnetic field strengths studied. Data acquisition and parameters are described in *Table III-VA*.

Table III-V
Backbone amide ^{15}N relaxation measurements

A. Sample ^{15}N-cCTnC free and ^{15}N-cCTnC complexed with ^{1}H-cIp

Experiment Name	^{1}H freq[a]	nt[b]	*x*-pts[c]	*y*-pts	*x*-sw[d]	*y*-sw
2D – ^{15}N-T_1[e]	500 / 600	16	768 / 896	96 / 96	6000 / 7200	1650 / 1800
2D – ^{15}N-T_2[f]	500 / 600	16	768 / 896	96 / 96	6000 / 7200	1650 / 1800
2D – $\{^1H\}$-^{15}N NOE	500 / 600	64	768 / 896	128 / 128	6000 / 7200	1650 / 1800

B. Sample ^{15}N-cIp free and ^{15}N-cIp complexed with cCTnC

Experiment Name	^{1}H freq[a]	nt[b]	*x*-pts[c]	*y*-pts	*x*-sw[d]	*y*-sw
2D – ^{15}N-T_1[e]	500 / 600	16	878 / 896	96 / 96	6000 / 7200	1650 / 1800
2D – ^{15}N-T_2[f]	500 / 600	16	878 / 896	96 / 96	6000 / 7200	1650 / 1800
2D – $\{^1H\}$-^{15}N NOE	500 / 600	64	878 / 896	96 / 96	6000 / 7200	1650 / 1800

[a] The frequency of ^{1}H in MHz.
[b] The number of transients acquired for each FID
[c,d] x,y-pts and sw is the number of complex points and sweep width in each respective dimension (*x* is the directly detected dimension)
[e] T_1 relaxation delays were 11.1, 55.5, 122.1, 199.8, 277.5, 388.5, 499.5, 666, 888 and 1100 ms on both instruments.
[f] T_2 relaxation delays were set to 16.61 **x** n and 16.544 **x** n ms (where n = 1,….., 11) on the 500 and 600 MHz spectrometers, respectively.

The relaxation values for the 8.2 kDa domain of ^{15}N-cCTnC•2Ca^{2+} upon the addition of the 2.5 kDa cIp peptide are consistent with calculated values for a 10.7 kDa binary complex. There is an increase for T_1 values and a decrease for T_2 values, with NOE values remaining largely unchanged when compared to cCTnC•2Ca^{2+}. The linker region and the H-helix of cCTnC display the largest changes in relaxation values upon cIp binding, indicating that this region is becoming more ordered upon cIp binding.

These results correspond well with the structure of the binary complex of ^{15}N-cCTnC•2Ca^{2+}•cIp, as multiple NOE contacts are observed between cIp and cCTnC•2Ca^{2+} in the linker region and the H-helix.

III-1. Backbone amide ^{15}N-NMR relaxation studies of ^{15}N-cIp

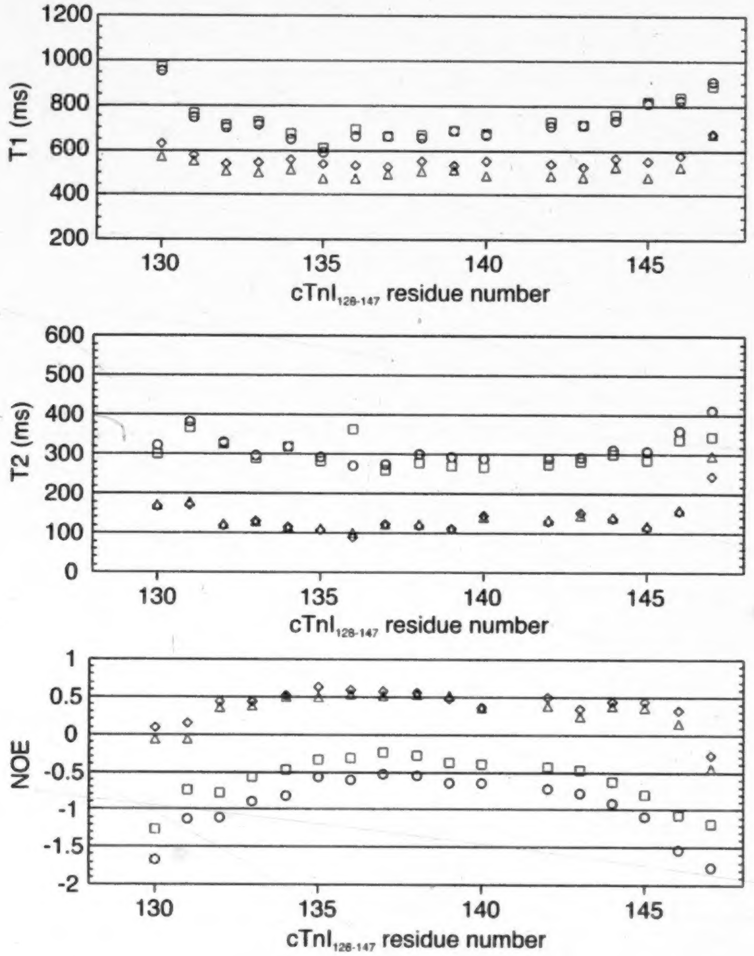


Figure III-12. **C-lobe ^{15}N-amide relaxation measurements.** ^{15}N-NMR relaxation values of ^{15}N-cIp unbound at 500 MHz (○), ^{15}N-cIp unbound at 600 MHz (□), ^{15}N-cIp bound to cCTnC•2Ca^{2+} at 500 MHz (△) and ^{15}N-cIp bound to cCTnC•2Ca^{2+} at 600 MHz (◊). T_1, T_2 values are reported in milliseconds (ms), and NOE values are a ratio of the intensity in the absence/presence of proton pre-saturation.

Backbone amide ^{15}N-NMR relaxation data for ^{15}N-cIp and cCTnC•2Ca^{2+}•^{15}N-cIp were obtained at 500 and 600 MHz (see Figure III-12). Backbone resonances for Thr128 and Gln129 were not observed due to rapid amide proton exchange with water for both samples. Residue Pro141 was not observed due to the absence of a residual amide proton. The T_1, T_2, and NOE values observed at the two frequencies are typical for an unstructured 2.5 kDa domain in solution. Addition of cCTnC•2Ca^{2+} produces large effects on the relaxation rates of all residues present on cIp. As expected for an increase in total molecular weight to 10.7 kDa (assuming a 1:1 ratio), decreases of both T_1 and T_2 values are observed at both magnetic field strengths, along with a large increase in the recorded values for NOE. Interpretation of the data reveals that cIp is becoming increasingly rigid upon cCTnC binding. Data acquisition and parameters are described in *Table III-VB*.

Comparison of the total molecular weights of both ^{15}N-cCTnC•2Ca^{2+}•cIp and cCTnC•2Ca^{2+}•^{15}N-cIp indicates that both systems should produce equivalent experimental correlation times (τ_C), assuming isotropic tumbling. Predicted τ_C values for a 1:1 binary complex of ^{15}N-cCTnC•2Ca^{2+}•cIp and cCTnC•2Ca^{2+}•^{15}N-cIp are 5.5ns and 5.5ns respectively, whereas experimentally measured values of 5.8ns and 5.7ns are observed, enforcing that the binary complex of cCTnC and cIp are binding in a 1:1 ratio. The measured T_2 values of ^{15}N-cCTnC•2Ca^{2+}•cIp mirror the values measured for cCTnC•2Ca^{2+}•^{15}N-cIp (~120ms) also supporting a 1:1 binary complex, as well implying that cIp is becoming increasingly rigid upon cCTnC binding even though the structure of cIp in the binary complex is not as well defined as cCTnC.

DISCUSSION

My NMR solution structure of the inhibitory region bound to the C-lobe required somewhat elaborate methods in the arrival of the completed structural ensemble. Generally when solving the structure of a peptide bound to a larger domain, the use of NMR experiments such as edited-filtered which take advantage of ^{13}C and/or ^{15}N magnetization transfers, are more than sufficient in determining inter-domain NOE's and generating restraints for structural calculations. However in this study, these edited-filtered experiments yielded very limited signal and thus usable information. Hence, the production of $^{13}C/^{15}N$-cIp peptide was the logical choice in the completion of a binary complex structure, as additional structural restraints arising from labeled inhibitory region could be added to the annealing protocol. As well, I was able to assign unambiguously all inter-domain cross-peaks in the ^{13}C-NOESY acquired for the $^{13}C/^{15}N$-cCTnC•$^{13}C/^{15}N$-cIp sample. The earlier edited-filtered work of the unlabeled peptide was instrumental in aiding in the initial placement of the peptide on the C-lobe, due to the unique chemical shifts of the Hβ protons of Phe132 and Phe138 in the complex (refer to Figure III-2), and was used as a starting point for all other inter-domain NOE contacts. During initial assignments for the ^{13}C-NOESY, an initial structural calculation was performed using only the Phe132 and Phe138 contacts to the C-lobe, which aided in the placement of the peptide on the C-lobe, which was then entered back into the nmrView assignment module to allow for the prediction of inter-domain NOEs with greater certainty. In total, 267 additional intra-peptide NOE restraints were also added to the annealing protocol for cIp that aided in obtaining a lower r.m.s.d. of cIp in the presented ensemble of 30 calculated structures.

The structure and dynamics of the inhibitory region of cardiac troponin I with the C-lobe of troponin C requires a detailed discussion regarding the previous 15+ years of prior structural/modeling data of this system. The general mechanism of contraction involving temporal interactions of the thin and thick filaments in muscle fibers upon a Ca^{2+} signaling cascade is now quite well understood (*73-77*), however when this project was started in 2000 there was very little known of the specifics of these interactions at the atomic level.

In the binary complex, cIp adopts a helical conformation, making NOE contacts with the linker region of cCTnC, as well as with both the E and H helices. Residues Leu134-Kys139 adopts a helical arrangement within the binary complex, with residues Arg140-Arg147 displaying an extended conformation, void of any secondary structure elements. Reflection on the structure of the binary complex reveals that all previous predicted models are in part incorrect. The inhibitory region has no β-sheet secondary structure present, as is indicated by the presence of traditional helical NOE's (i.e. $d_{\alpha\beta(i,\,i+3)}$, $d_{\alpha N(i,\,i+3)}$), and by chemical shift data for $^1H_\alpha$ and $^{13}C_\alpha$ atoms in the protein backbone. No long range NOE ($|i-j| \geq 5$) contacts are observed, which would be predicted if the inhibitory region possessed β-sheet characteristics. It is easy to comprehend why Hernandez made the prediction of an extended inhibitory region (see Figure III-13) as they limited their study to only ^{1}H-NMR chemical shifts changes, which only deviated moderately from what one would expected from an extended conformation (*28*). I also observed the same moderate changes in the ^{1}H-NMR chemical shifts (towards a helical structure) upon complex formation, however the changes observed in the ^{13}C-C_α chemical shifts were without question shifted more towards a helical secondary structure in the binary complex.

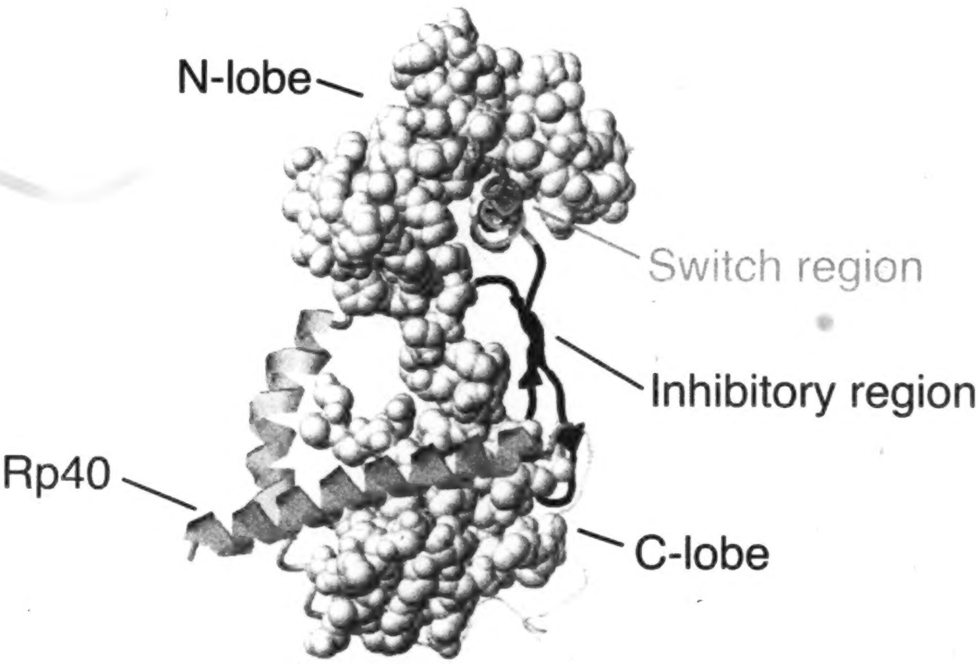


Figure III-13. **Model of *skeletal* troponin I in complex with troponin C.** The proposed model by the Trewhella group places the inhibitory region as an anti-parallel β-sheet facing (red) away from the hydrophobic face of the C-lobe. Troponin I domains are colored labeled for clarification. Adapted from (*38*).

The subsequent model proposed by the Trewhella group (*38*) was based, in part, on the skeletal ^{1}H-NMR work performed by Hernandez, and thus it should be asked if it is indeed valid to compare the skeletal results to that of my cardiac studies. Since the inhibitory region for both the skeletal and cardiac isoforms differ in only one position (Thr142 in cardiac is a Pro in skeletal), and since troponin C from both skeletal and cardiac isoforms is well conserved, it is immediately applicable to make a direct comparison of the skeletal model to my presented structure.

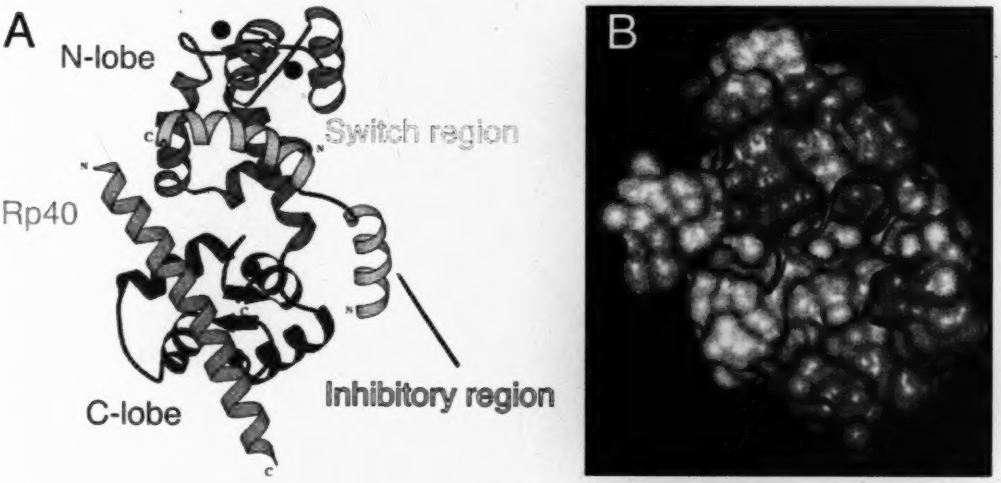


Figure III-14. **Helical models of the inhibitory region of skeletal troponin I.** (A) Vassylyev model places the inhibitory region away from the C-lobe. (B) Ngai model places the inhibitory region in the hydrophobic face of the C-lobe with helical turns on either side of the central Pro-Pro bond. Panels A and B were taken from (*18*) and (*19*).

The helical propensity of the inhibitory region was correctly modeled separately by both Vassylyev and Ngai (see Figure III-14), however the positioning of the helix in reference to the C-lobe was incorrect in both cases (*30, 33, 78*). Ngai originally modeled the inhibitory region (sTnI$_{104-115}$) of the skeletal isoform, using cross-linking techniques and previous NMR data based on transferred NOE experiments by Campbell, where the peptide adopts an amphiphilic helix-like structure (*30, 78*), to place it in the hydrophobic core of the C-lobe. However this model incorrectly placed helical propensity both before and after the Pro-Pro bond in a 'collapsed' model. This 'collapsed' effect is due to the modeling of the C-terminal tail of the inhibitory peptide adopting a helical conformation instead of being extended and pointing out of the hydrophobic core. In the Vassylyev model, the helical propensity of the inhibitory region was placed away from the hydrophobic face, in a position that occupied the back face of the C-lobe.

My binary structure closely matches the methionine-spin label NMR study reported by Howarth, which surface exposure protection for only Met120 and Met157 of the C-lobe during cIp (Asn129-Ile148) association was observed (*29*). These residues are displayed in Figure III-15, revealing how Met120 and Met157 become structurally buried upon cIp binding. Met103, Met120, Met137 and Met157 are all solvent exposed in the C-lobe when unbound, however when cIp binds residues Met120 and Met157 become protected from solvent effects.

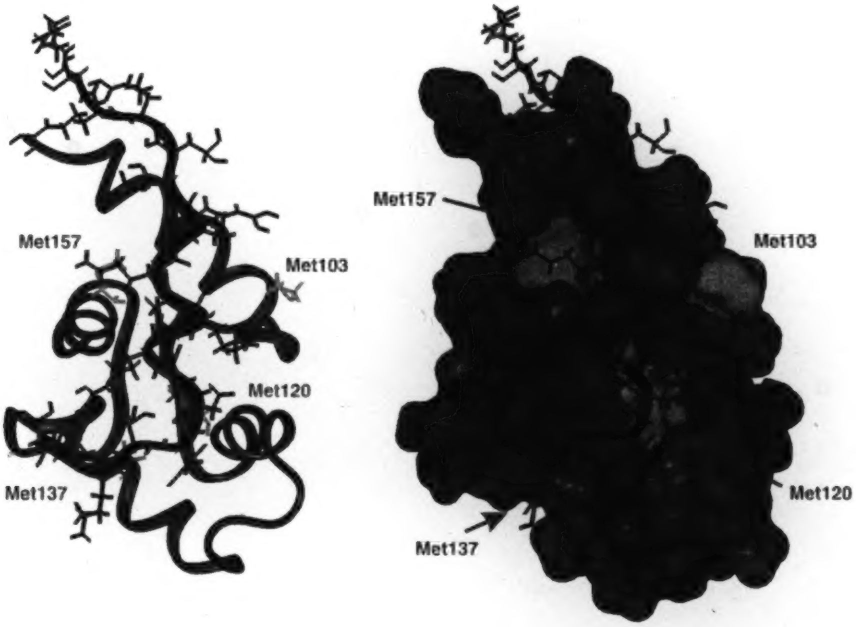


Figure III-15. **Methionine solvent protection of C-lobe upon cIp binding.** Ribbon diagram (left) and Connolly surface representation (right) of a single structure from the binary complex ensemble with residues Met103, Met120, Met137 and Met157 diagramed in green. Met120 and Met157 become solvent inaccessible upon the binding of cIp, however Met103 (positioned on the back-face) and Met137 remain solvent accessible.

Dr. Michael L. Love from the Cohen laboratory in Brandeis University reported a crystal structure of a troponin I peptide (Asn96-Lys123) bound to troponin C (*79*) in his May 2000 Ph.D. thesis (unpublished structure). This structure dealt exclusively with the skeletal isoform. Although never published, residues Leu111-Asp119 had electron density (corresponding to residues 143-151 of the cardiac isoform) and were found bound exclusively to the N-lobe with contacts to the C and D helices. An in depth analysis of the structure is unavailable as the pdb coordinates were never deposited. However, based upon his figures, this structure does not correspond well with my presented structure and places the inhibitory region far away from the C-lobe. As well it does not correspond with the structure of the switch peptide bound to the N-lobe within the cardiac/skeletal isoforms (*52, 80, 81*). The switch region is normally helical from residues 118-124 (skeletal numbering), however it was not visualized within this complex. Not surprisingly, this structure was never published as it did not fit well with the majority of current biochemical and biophysical understandings, revealing that this crystal structure might include induced non-specific contacts that might be occurring due to the low temperatures involved in crystallography methods.

I have highlighted and discussed some of the more significant structural models and ideas regarding the precise placement of the inhibitory region of troponin I, I have also attempted to make a model of the troponin complex in context with my presented inhibitory structure (*82*). Note that this modeling was performed in early 2003 based upon available data and was presented prior to the published structure of the 'core' troponin complex by Takeda (*39*). It has been well established that the RP40 region in the skeletal isoform will displace the inhibitory region off the C-lobe due to an increased binding affinity of the RP40 region (*19*). It is due to this high affinity (<1*u*M), that it is concluded that the RP40 region is always bound to the C-lobe of TnC, and does not move on-and-off the thin filament during the calcium-signaling cascade like other regions (i.e. cSp) of TnI. When considering my data regarding the C-lobe, a model can be presented as to how the two can exist simultaneously on the C-lobe that explains the skeletal results for RP40 competition binding. I have modeled the inhibitory region on the C-lobe, crossing over cRP40 on the hydrophobic face in Figure III-16 (*82*). The positioning of cRP40 is consistent with the crystal structure by Vassylyev of the skeletal $sCTnC_{1-47}$•sRP40 structure. In addition, I have placed the switch peptide bound to the N-lobe as per the structure by Li and Spyracopoulos (*52*), and have placed the N- and C-lobes in an orientation as per Dvoretsky (*36*). Although the cRP40 region will displace and move the inhibitory region off the hydrophobic face of the C-lobe, the inhibitory region remains bound overtop of the cRP40 region due to suggested electrostatic interactions with the C-lobe along the E-helix. This model also places part of the inhibitory region in close vicinity to the D/E linker, which agrees with prior results that the two may have direct interactions with one another.

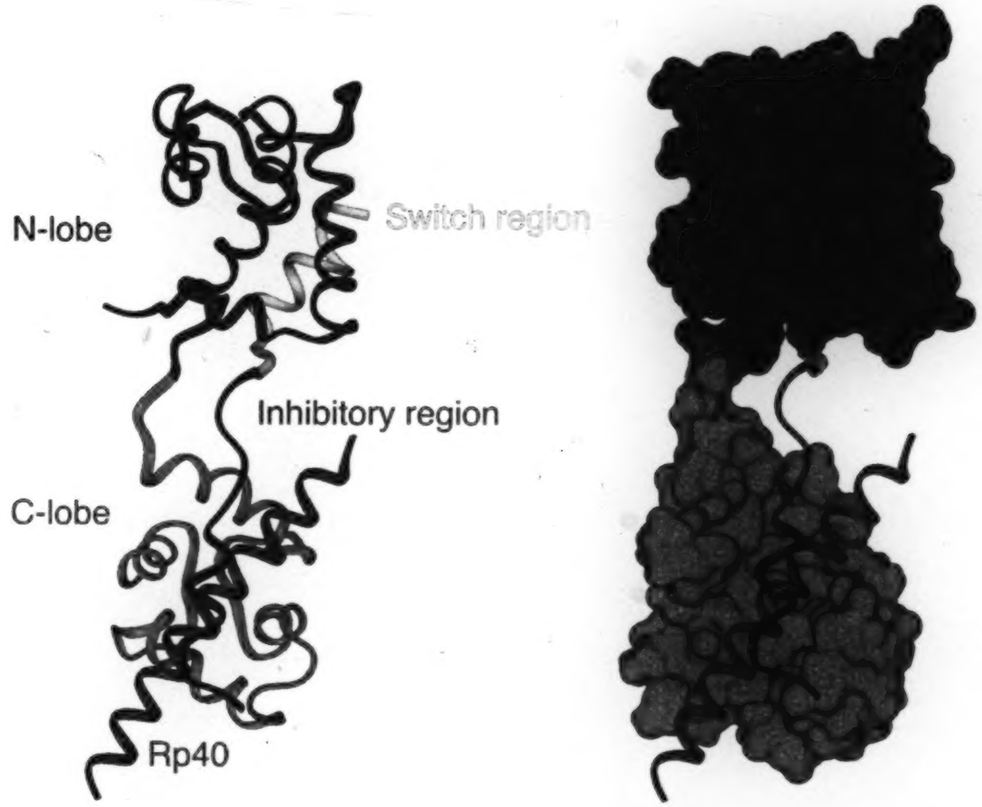


Figure III-16. **Lindhout model of inhibitory region bound to cardiac troponin C**. A ribbon diagram model is shown to the left and a ribbon and connolly cTnC-surface of is shown to the right. Individual domains are color indicated for clarity.

Within the binary complex presented here, numerous stabilizing electrostatic interactions occur between the acidic C-lobe and the highly basic inhibitory region. Electrostatic NOE contacts are difficult to measure by NMR, due to proton exchange and NOE distance limitations (<5Å). The probability of stabilizing electrostatic interactions within the complex is high, as inspection of the structure yields close association of basic/acidic residues (see Figure III-10), with the potential of the unstructured region containing residues Arg140, Arg144, Arg145 and Arg147 in cIp to interact with the acidic residues Glu94, Glu95 and Glu96 present on the E-helix. Interestingly, recent work by Tripet (*35*) has shown a decrease in the affinity of cIp to cTnC with increasing concentrations of KCl, re-enforcing the idea that electrostatics may play a part in binding. The potentials of electrostatic interactions pertaining to this complex are explored and discussed in Chapter IV of this thesis using mutational and phosphorylation events, and therefore will not be extensively discussed within this chapter.

With the availability of large amounts of recombinant peptide, the potential to measure backbone restraints via ^{15}N-NMR RDCs for the bound inhibitory region was proposed, as this methodology would allow for an additional structural restraint. Based upon the relaxation data, I was able to conclude that the C-terminal residues Arg140-Arg147 of the inhibitory region were bound tightly to the C-lobe, however I was unable to define any contacts between this section of the inhibitory region with the C-lobe. The measurement of RDCs might alleviate the uncertainty of the placement of the C-terminal tail. In a collaborative effort with Grant L. Gay in the Sykes laboratory, Grant measured RDCs of the bound peptide to cTnC using lanthanide ions bound to the metal binding sites of cTnC, to establish partial alignment of the protein complex (see Appendix E). Unfortunately he achieved very limited alignment and was only able to measure RDCs on the order of a few Hz for the helical region of cIp. However the limited RDC values that Grant acquired closely matched the secondary structural elements that I presented in my binary complex. The use of lanthanide ions to establish alignment was perhaps not the best of options for obtaining RDCs and the use of newer methods such as poly-acrylamide stretched gels methodology might have yielded stronger RDCs for structural calculations (*83-86*), however this project was never completed.

The previously mentioned models and structures (including my own structure and model) all dealt with fragments of TnI and/or TnT in complex with TnC. Results obtained from groups which have utilized the full length troponin complex must also be addressed. Recent results published by Brown addressed the structure of the inhibitory region by spin labeling EPR (*34*). Within this study it was proposed that the inhibitory region in a ternary troponin complex possessed a helical region from residues Gln129-Lys137, with residues Phe138-Arg145 showing no secondary structure elements. The helical spanning region of the inhibitory region was shown to consist of a coiled-coil motif, where Lys130-Arg135 of cIp are in contact with TnT. These results correlated rather well with my presented binary structure, specifically the helical spanning region that I reported; Leu134 to Lys139. Due to these results, it is not surprising that the helical fidelity of the inhibitory region in my binary complex was compromised, most likely due to the fact that a TnT binding partner was not present to stabilize the helical fold. Therefore the previously mentioned amphipathic helix proposed for the inhibitory region by Campbell was in part correct, however it was modeled incorrectly by Ngai (*30, 33, 78*).

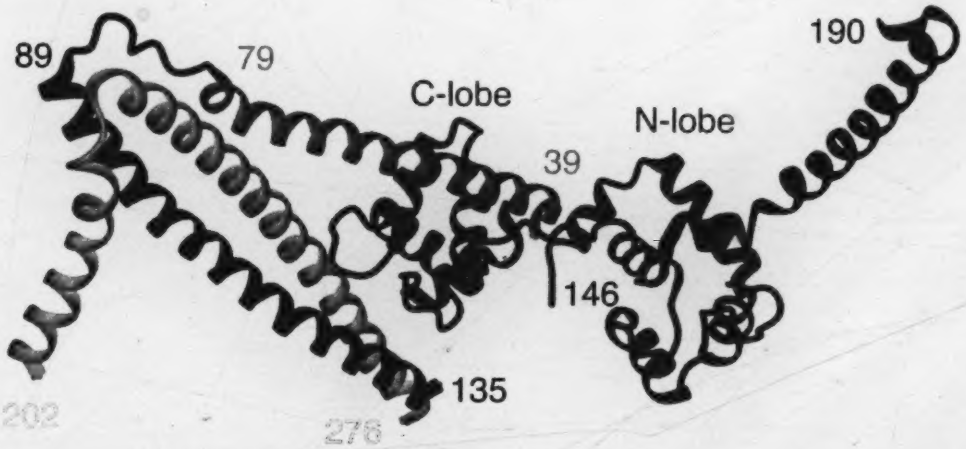


Figure III-17. **Crystal structure of the 'core' cardiac troponin 52kDa complex**. Ribbon diagram of 1J1E.pdb where cTnC lobes are colored in red, cTnI regions are colored in blue and cTnT regions are colored in yellow. Residue number is included for reference.

The July 3rd, 2003 publication of the crystal structure of a 'core' troponin complex by Takeda presented a structure different from all previously measured models. In the published structure, Takeda presented two different structures in the calcium saturated state, one of a 46kDa complex corresponding to a cTnC•cTnI$_{31\text{-}163}$•cTnT$_{183\text{-}288}$ complex, and a second 52kDa complex consisting cTnC•cTnI$_{31\text{-}210}$•cTnT$_{183\text{-}288}$, which contained a helical extension of cTnI beyond the C-terminal switch region. These crystal structures displayed the first high-resolution structure(s) of the complete troponin complex (see Figure III-17). The predicted coiled-coil motif between the T and I subunits was now detailed in atomic resolution; spanning five heptad repeats within the coiled-coil and making contacts to the C-lobe of cTnC. The 'IT' arm, as the authors described it, was a novel structure and showed the I-T interaction responsible for holding the core troponin complex to the thin filament via troponin T. In addition, the structure validated previous years of work by numerous structural groups such as the Li and Spyracopoulos NMR structure of the N-lobe in complex with the switch region (*52*). These crystal structures also provided a validation to the past years of previously NMR and other biophysical studies that broke the complex up into smaller fragments. *It appears that the parts of the sum indeed do behave, for the most part, in an identical manner to those of the whole.*

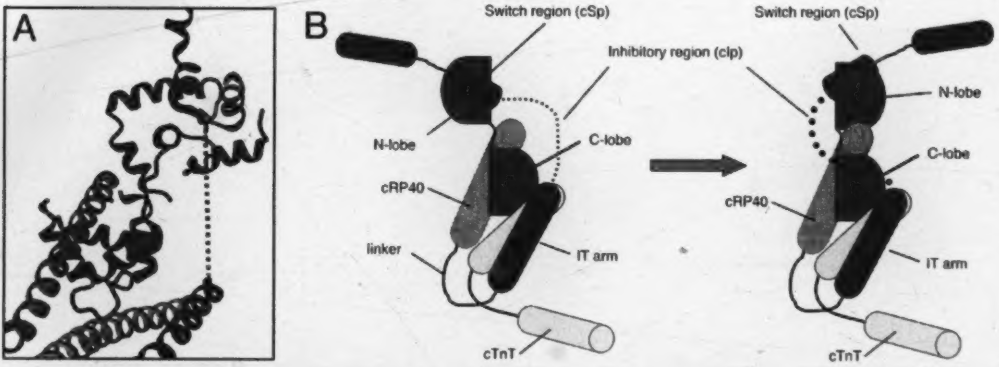


Figure III-18. **Model of the 'core' cardiac troponin complex of the N-lobe.** Hypothetical re-orientation of the N-lobe of cTnC around the D/E helical region. (A) Ribbon diagram of the crystal structure 1J1E with a dotted line displaying an extended inhibitory region, away from the C-lobe. (B) Cartoon model of the Takeda structure (left) that is shown in the same relative orientation as in panel A, where a rotation of the N-lobe (right) would allow the inhibitory region to come into contact with the C-lobe.

In regards to my structure of the binary complex, the disappointing feature of the 'core' crystal structure is the lack of electron density in the inhibitory region spanning from residues Arg135-Val146 (see Figure III-18A). In the crystal structure, Takeda and Maeda (1J1D.pdb) predicted this region to be fully extended and that it must span greater than 20Å to connect the two ends of the polypeptide chain. In part, my binary structure agrees with Takeda that the inhibitory region is extended, as I have shown residues Arg140-Arg147 are indeed void of secondary structure. However the implied placement of the inhibitory region in the crystal structure is in direct conflict with my binary structure. In the Takeda structure, the distances from the region connecting Arg135 to Val146 make it impossible for close contacts with the C-lobe (see Figure III-17). Therefore how does my relaxation data, which revealed that the inhibitory residues Arg140-Arg147 are bound tightly to the C-lobe, fit with the crystal structure? A few explanations are proposed to address this discrepancy. One possibility is the positioning of the N-lobe. In the Takeda structure, there are virtually no restraints that hold the N-lobe in a fixed position. Perhaps in the crystal structure, Takeda observed an unusual orientation of the N-lobe? The orientation of the N-lobe relative to the C-lobe is in sharp contrast with the lobe orientation presented by Dvoretsky (see Figure III-17 model). Given the potential for flexibility of the N-lobe orientation, a movement of the N-lobe around the D/E linker of cTnC, which would bring the inhibitory region in a closer proximity to the C-lobe, would more closely parallel my binary structural results (see Figure III-18B). Due to the inherently low temperatures required for crystal diffraction, the domains may have been locked into a static, low-temperature orientation that may not have been a true representation of the 'flexible' cardiac Tn complex at physiological conditions and temperatures. A second possibility is that while my binary complex is correct, I did not have all the troponin binding partners present. Thus the two structures can individually stand-alone and do not require a direct comparison between the two as it would not be applicable. A third possibility is that when the cRP40 region is bound the C-lobe, it completely abolishes the binding affinity of the inhibitory region for the C-lobe and thus the binary complex is not relevant, and the crystal structure is correct that the inhibitory region is away from cTnC. A fourth possibility is that both the crystal structure and my NMR binary structure are incorrect in part, as the conformations and orientations of the domains might be altered when incorporated within of the thin filament *in vivo*.

Based upon the discrepancies of the binary complex to that of the solved crystal structure, it becomes inherently difficult to predict physiological roles of the inhibitory region based solely on the binary complex. Further studies on the role of the binary complex must be performed. An in-depth analysis of the electrostatic potentials of the inhibitory region must be pursued prior to an explanation regarding events such FHC mutations and regulation by phosphorylation is attempted. As well, what role does the inhibitory region play with interactions with the N-lobe? Does the cardiac isoform display the same competition between the inhibitory region and cRP40 for the C-lobe? Does the cardiac 'core' troponin complex behave identical in solution as we observe in the crystal structure, and are there areas of intrinsic flexibility that we can monitor of the 'core' complex? I have addressed these issues in Chapters IV, V, VI and VII of this thesis in an attempt to ascertain an understanding of the inhibitory region of cTnI in solution. Within these chapters, a discussion regarding relevant physiological and biophysical data will be expanded. Chapter VIII will present a comprehensive view of cardiac contraction by piecing together all the relevant structural and biophysical data from not only my own work, but by those in the muscle field who have made contributions in the understanding of the structural and physiological events occurring in contraction. Additionally, comparisons of the cardiac crystal structure to that of a recently published crystal structure (March 2005) of the skeletal troponin isoform will be addressed in Chapters V, VI and VIII of this thesis (*87*).

REFERENCES

1. Herzberg, O., and James, M. N. G. (1985) *Nature 313*, 653-659.
2. Herzberg, O., and James, M. N. G. (1988) *J. Mol. Biol. 203*, 761-779.
3. Kawasaki, H., and Kretsinger, R. H. (1994) *Protein Profile 1*, 343-517.
4. Kawasaki, H., and Kretsinger, R. H. (1995) *Protein Profile 2*, 297-490.
5. Sia, S. K., Li, M. X., Spyracopoulos, L., Gagné, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) *J. Biol. Chem. 272*, 18216-18221.
6. Luo, Y., Wu, J.-L., Gergely, J., and Tao, T. (1997) *Biochemistry 36*, 11027-11035.
7. Filatov, V. L., Katrukha, A. G., Bulargina, T. V., and Gusev, N. B. (1999) *Biochemistry (Mosc). 64*, 969-85.
8. Tao, T., Gong, B. J., Grabarek, Z., and Gergely, J. (1999) *Biochim Biophys Acta 1450*, 423-33.
9. Maytum, R., Geeves, M. A., and Lehrer, S. S. (2002) *J Biol Chem 277*, 29774-80.
10. Perry, S. V. (1999) *Mol. Cell. Biochem. 190*, 9-32.
11. Solaro, R. J., el-Saleh, S. C., and Kentish, J. C. (1989) *Mol Cell Biochem 89*, 163-7.
12. Head, J. F., and Perry, S. V. (1974) *Biochemical Journal 137*, 145-154.
13. Talbot, J. A., and Hodges, R. S. (1979) *J. Biol. Chem. 254*, 3720-3723.
14. Syska, H., Wilkinson, J. M., Grand, R. J., and Perry, S. V. (1976) *Biochem J 153*, 375-87.
15. Li, M. X., Spyracopoulos, L., Beier, N., Putkey, J. A., and Sykes, B. D. (2000) *Biochemistry 39*, 8782-8790.
16. Ngai, S.-M., and Hodges, R. S. (1992) *J. Biol. Chem. 267*, 15715-15720.
17. Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) *J Mol Biol 271*, 728-50.
18. Vassylyev, D. G., Takeda, S., Wakatsuki, S., Maeda, K., and Maeda, Y. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 4847-4852.
19. Mercier, P., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 2902-11.
20. Solaro, R. J. (1999) *Circ Res 84*, 122-4.
21. Solaro, R. J. (2003) *Adv Exp Med Biol 538*, 389-401; discussion 401-2.
22. Noland, T. A., Jr., Guo, X., Raynor, R. L., Jideama, N. M., Averyhart-Fullard, V., Solaro, R. J., and Kuo, J. F. (1995) *J Biol Chem 270*, 25445-54.
23. Noland, T. A., Jr., Raynor, R. L., Jideama, N. M., Guo, X., Kazanietz, M. G., Blumberg, P. M., Solaro, R. J., and Kuo, J. F. (1996) *Biochemistry 35*, 14923-31.
24. Chandra, M., Dong, W. J., Pan, B. S., Cheung, H. C., and Solaro, R. J. (1997) *Biochemistry 36*, 13305-11.
25. Deng, Y., Schmidtmann, A., Redlich, A., Westerdorf, B., Jaquet, K., and Thieleczek, R. (2001) *Biochemistry 40*, 14593-602.
26. Kranias, E. G., and Solaro, R. J. (1982) *Nature 298*, 182-4.
27. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.
28. Hernandez, G., Blumenthal, D. K., Kennedy, M. A., Unkefer, C. J., and Trewhella, J. (1999) *Biochemistry 38*, 6911-7.

29. Howarth, J. W., Krudy, G. A., Lin, X., Putkey, J. A., and Rosevear, P. R. (1995) *Protein Sci. 4*, 671-680.
30. Campbell, A. P., and Sykes, B. D. (1991) *J Mol Biol 222*, 405-21.
31. Campbell, A. P., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) *Biochim Biophys Acta 1160*, 35-54.
32. Mercier, P., Spyracopoulos, L., and Sykes, B. D. (2001) *Biochemistry 40*, 10063-77.
33. Ngai, S.-M., Sönnichsen, F. D., and Hodges, R. S. (1994) *J. Biol. Chem. 269*, 2165-2172.
34. Brown, L. J., Sale, K. L., Hills, R., Rouviere, C., Song, L., Zhang, X., and Fajer, P. G. (2002) *Proc Natl Acad Sci U S A 99*, 12765-70.
35. Tripet, B., De Crescenzo, G., Grothe, S., O'Connor-McCourt, M., and Hodges, R. (2002) *J Mol Biol 323*, 345.
36. Dvoretsky, A., Abusamhadneh, E. M., Howarth, J. W., and Rosevear, P. R. (2002) *J Biol Chem 277*, 38565-70.
37. Heller, W. T., Abusamhadneh, E., Finley, N., Rosevear, P. R., and Trewhella, J. (2002) *Biochemistry 41*, 15654-63.
38. Tung, C. S., Wall, M. E., Gallagher, S. C., and Trewhella, J. (2000) *Protein Sci 9*, 1312-26.
39. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
40. Putkey, J. A., and Liu, W. (2002) *Methods Mol Biol 172*, 151-60.
41. Li, M. X., Gagné, S. M., Tsuda, S., Kay, C. M., Smillie, L. B., and Sykes, B. D. (1995) *Biochemistry 34*, 8330-8340.
42. Huth, J. R., Bewley, C. A., Jackson, B. M., Hinnebusch, A. G., Clore, G. M., and Gronenborn, A. M. (1997) *Prot. Sci. 6*, 2359-2364.
43. Lindhout, D. A., Thiessen, A., Schieve, D., and Sykes, B. D. (2003) *Protein Sci 12*, 1786-91.
44. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
45. Johnson, A. J., and Blevins, R. A. (1994) *Journal of Biomolecular NMR 4*, 603-614.
46. Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry 36*, 4386-4392.
47. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thorton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-290.
48. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr 54 (Pt 5)*, 905-21.
49. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984-6003.
50. Boyko, R., and Sykes, B. D. (2003) *XCRVFIT, Protein Engineering Networks of Centres of Excellence v.2.1.1*, http://www.pence.ca/software/xcrvfit/index.html.
51. Wang, X., Li, M. X., and Sykes, B. D. (2002) *J Biol Chem 277*, 31124-33.
52. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.

53. Wang, X., Li, M. X., Spyracopoulos, L., Beier, N., Chandra, M., Solaro, R. J., and Sykes, B. D. (2001) *J Biol Chem 276*, 25456-66.
54. Ogura, K., Terasawa, H., and Inagaki, F. (1996) *J. Biomol. NMR 8*, 492-498.
55. Neri, D., Szyperski, T., Ottig, G., Senn, H., and Wuthrich, K. (1989) *Biochemistry 28*, 7510-7516.
56. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.
57. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.
58. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR 4*, 871-878.
59. Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson. B103*, 203-216.
60. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) *J. Magn. Reson. B101*, 333-337.
61. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1994) *J. Magn. Reson. B103*, 197-201.
62. Lindhout, D. A., and Sykes, B. D. (2003) *J Biol Chem 278*, 27024-34.
63. Spyracopoulos, L., Li, M. X., Sia, S. K., Gagné, S. M., Chandra, M., Solaro, R. J., and Sykes, B. D. (1997) *Biochemistry 36*, 12138-12146.
64. Slupsky, C. M., and Sykes, B. D. (1995) *Biochemistry 34*, 15953-15964.
65. Cachia, P. J., Sykes, B. D., and Hodges, R. S. (1983) *Biochemistry 22*, 4145-52.
66. Mercier, P., Ferguson, R.E., Irving, M., Corrie, J.E.T., Trentham, D.R. and Sykes, B.D. (2003) *Biochemistry submitted.*
67. Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol. 239*, 363-392.
68. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171-180.
69. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311-333.
70. Wishart, D. S., Boyko, R. F., Willard, L., Richards, F. M., and Sykes, B. D. (1994) *Comput Appl Biosci 10*, 121-132.
71. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol 277*, 505-524.
72. Nicholls, A. (1992) *GRASP: graphical representation and analysis of surface properties.*, Columbia University, New York.
73. Adelstein, R. S., and Eisenberg, E. (1980) *Annu Rev Biochem 49*, 921-56.
74. Leavis, P. C., and Gergely, J. (1984) *CRC Crit. Rev. Biochem. 16*, 235-305.
75. Zot, A. S., and Potter, J. D. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 535-59.
76. Tobacman, L. S. (1996) *Annu. Rev. Physiol. 58*, 447-481.
77. Farah, C. S., and Reinach, F. C. (1995) *FASEB J. 9*, 755-767.
78. Campbell, A. P., and Sykes, B. D. (1989) *Adv Exp Med Biol 255*, 195-204.
79. Love, M. L. (2000) *Thesis of Dr. Michael L. Love*, http://proclus.tripod.com/thesis/index.html.
80. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1998) *Biophys. J. 74*, A51.
81. Mercier, P., Ferguson, R. E., Irving, M., Corrie, J. E., Trentham, D. R., and Sykes, B. D. (2003) *Biochemistry 42*, 4333-48.
82. Sykes, B. D. (2003) *Nat Struct Biol 10*, 588-9.

83. Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M., and Bax, A. (1997) *Nat Struct Biol 4*, 732-8.
84. Tjandra, N., and Bax, A. (1997) *Science 278*, 1111-4.
85. Bax, A., and Tjandra, N. (1997) *J Biomol NMR 10*, 289-92.
86. Ottiger, M., Delaglio, F., Marquardt, J. L., Tjandra, N., and Bax, A. (1998) *J Magn Reson 134*, 365-9.
87. Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) *Proc Natl Acad Sci U S A*.

CHAPTER IV

The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C

OVERVIEW

This chapter describes studies designed to elucidate the nature of the interactions revealed in the NMR solution structure of the C-lobe of cTnC in complex with the inhibitory region of cTnI, which was solved and reported in Chapter III of this thesis. In Chapter III, electrostatic interactions occurring between the highly basic inhibitory region and the acidic C-lobe were proposed as the primary force responsible for domain association. To test this hypothesis, I have undertaken a Glycine scanning mutagenesis study of the inhibitory region and have monitored binding changes via NMR spectroscopy. In addition to glycine mutations, I have looked at a familial hypertrophic cardiomyopathy mutation, which occurs within the inhibitory region that is linked to a diseased state of cardiac tissue (R144G), and have looked at the phosphorylation of Thr142 which is a known protein kinase C phosphorylation site that plays a direct role in myofibril regulation. Conclusions based upon this chapter of work supports the theory that electrostatic attractions are the dominant force acting between these two domains. In addition, the regulation of the thin filament by phosphorylation and consequences of mutation will be addressed in relation to the *in vitro* data reported here. All work presented in this chapter was performed solely by myself, with additional input from Xu Wang, Dean Schieve, Monica X. Li, Nina Buscemi, Jennifer E. Van Eyk, and Brian D. Sykes. A portion of the work presented in this chapter has been previously published in 2 articles submitted to the journal *Biochemistry*, specifically on June 11th, 2002 [*Biochemistry*. 2002 Jun 11;**41(23)**:7267-7] and on December 16th, 2003 [*Biochemistry*. 2003 Dec 16;**42(49)**:14460-8]. The remaining data in this chapter (unpublished) is currently in being prepared as a manuscript for publication to the journal *Biochemistry*, which should be submitted by July 2005.

INTRODUCTION

The myofilament is the contractile machinery in cardiac muscle cells and plays a vital role in maintaining the normal function of heart. The contractile proteins include myosin, actin, tropomyosin, and troponin. Troponin is a 1:1:1 complex of troponin C, troponin I, and troponin T. During diastole, troponin holds tropomyosin in a conformational state that blocks the interaction between myosin and actin. When Ca^{2+} binds to troponin C during systole, the troponin-tropomyosin complex changes so that it no longer inhibits the interaction between actin and myosin. This leads to tension producing cross bridges between actin and myosin, potentiating actomyosin ATPase activity, and ultimately heart muscle contraction (for recent reviews, see Refs (*1, 2*)).

Cardiac troponin I (cTnI) is the inhibitory component of the troponin complex and its interaction with cardiac troponin C (cTnC) plays a critical role in transmitting the Ca^{2+} signal to the other myofilament proteins in heart muscle contraction. The solution structure of the Ca^{2+}-saturated cTnC has revealed a dumbbell molecule with two globular domains connected by a flexible linker (*3*). The solution structures of the N-lobe of cTnC in both apo and Ca^{2+}-bound states have defined the Ca^{2+}-induced conformational transition in the regulatory domain of cTnC (*4*). Using the structure of cTnC as a framework, biochemical and biophysical studies have clarified how various regions of cTnI react with cTnC and participate in thin filament Ca^{2+}-signaling (for a review, see (*5*)). An antiparallel arrangement between cTnC and cTnI has been proposed and an inhibitory region has been identified (for a review, see (*5*)). The inhibitory region of cTnI is a 20 amino acid motif encompassing residues 128-147 in cTnI (cIp). This region is evolutionarily conserved and alternatively binds to either actin-tropomyosin or cTnC, depending on the intracellular concentration of Ca^{2+}. During diastole, when the $[Ca^{2+}]$ is low, the inhibitory region of cTnI is bound strongly to actin-tropomyosin, inhibiting the power stroke. During systole, transient increases in $[Ca^{2+}]$ promotes Ca^{2+} binding to cTnC, facilitating a shift of this region of cTnI from the thin filament to cTnC. Thus, a movement of the inhibitory region of cTnI from cTnC to actin-tropomyosin constitutes the major switch between muscle contraction and relaxation, and this switch is modulated by the interaction of the C-terminal domain of cTnI and cTnC (*6, 7*). Structural data on cTnI includes the early structure of the inhibitory peptide bound to cTnC determined by Campbell *et al.* (*8*) and additional structural information generated by using NMR and selective isotope labeling on Met residues in both cTnC and cTnI (*9-14*). In the NMR solution structure of the N-lobe of cTnC complexed to the switch region ($cTnI_{147-163}$ – cSp), this region, which immediately follows C-terminal to the inhibitory region of cTnI, binds to the hydrophobic pocket of the N-lobe of cTnC and stabilizes an open conformation (*15*). The NMR solution structure of the cardiac inhibitory region bound to the C-lobe of cTnC (see Chapter III) revealed a complex in which the inhibitory region spans the C-lobe making numerous contacts with the E- and H-helices and the F/G linker region (*16*). A resulting hypothesis of this structure was that the interactions between the C-lobe and the inhibitory region are predominantly electrostatic in nature. It is therefore not surprising that many known mutations that are know to cause disease, as well as troponin regulatory mechanisms (i.e. phosphorylation), are a direct consequence of changes within the electrostatic environment present on the thin filament.

cTnI has many unique characteristics. The human cTnI is a 210-residue protein and it differs from its skeletal isoform (sTnI) by a 33 amino acid N-terminal extension. There are also important sites of phosphorylation in cTnI that are not present in its skeletal counterpart. Two serine residues (serines 22 and 23), that are substrates for protein kinase A (PKA) lie within the N-terminal extension (for a review, see (*5*) and also (*12, 17*)). cTnI is also sensitive to protein kinase C (PKC) phosphorylation at Ser41, Ser43, and Thr142 (for a review, see (*5*)). The effect of cTnI phosphorylation is primarily a downward modulation of cardiac contractility, by decreased actomyosin ATPase activity, which is mediated by PKC. cTnI is also implicated in various cardiac diseases. For example, familial hypertrophic cardiomyopathy (FHC), one of the most frequently occurring inherited cardiac disorders, affecting up to 1 in 500 of the population and has been identified to be a sarcomeric disease (for reviews, see (*18, 19*)). It is characterized by left ventricular hypertrophy, myofibril disarray, sudden cardiac death, and is believed to be caused by mutations in certain contractile protein genes, including cTnI (i.e. R144G, R144Q, R161W, S198N, G202S, and K205Q). In addition to the reported mutations, a deletion of one codon (ΔK182) and a deletion of Exon8 (Δexon8) are also known areas which may compromise cTnI fidelity (*19*). Among those, R144G and R144Q are located in the key inhibitory region of cTnI. A truncation of cTnI has been associated with myocardial stunning (*20, 21*) and measuring the serum level of cTnI has become the standard of care in the diagnosis of myocardial injury (*22-25*).

In order to understand how alterations in cTnI cause disease, it is important to elucidate the effects of cTnI modifications on its interaction with cTnC. In this chapter, I have evaluated the effects of Thr142 phosphorylation and the FHC mutation R144G on the interaction of the important inhibitory region with the C-lobe. In addition to these known *in vivo* modifications, I have chosen to broaden the scope of this study to include mutations of other basic residues present on the inhibitory region (i.e. Lys139, Arg140, Arg145) to obtain a greater understanding of the nature of the electrostatic forces involved in domain association. A series of glycine-scanning mutagenesis of the basic residues (Arg or Lys) of the inhibitory region was chosen (as opposed to alanine-scanning mutagenesis) in an attempt to mirror the known biological mutation of the FHC mutation R144G.

For the purpose of this study, I have chosen an isolated Ca^{2+}-saturated C-lobe (residues 91-161) of human cTnC ($cCTnC•2Ca^{2+}$) since my earlier work has shown that the inhibitory peptide primarily interacts with the C-terminal domain of cTnC (*16, 26, 27*). The use of the smaller domain reduces the NMR spectral overlap and facilitates the assignment process. 2D $\{^1H, ^{15}N\}$-HSQC NMR spectroscopy was used to monitor the binding of eight differing inhibitory peptides [wild-type cIp, K139G, R140G, R144G, R145G, T142Phos, R144G-T142Phos and G4 (K139G, R140G, R144G, R145G)] respectively, to $cCTnC•2Ca^{2+}$. Chemical shift mapping indicates that while the binding epitope of cIp on $cCTnC•2Ca^{2+}$ is not greatly affected, the affinity can be reduced by up to 2 orders or magnitude from the wild-type cIp peptide (31*u*M) upon mutation. This suggests that even minor modifications of cIp can have an adverse effect on the binding of cIp to $cCTnC•2Ca^{2+}$. These perturbations may correlate with the impairment or loss of cTnI function in heart muscle contraction. All modifications will be directly addressed against the solved NMR solution structure of the C-lobe in complex with the inhibitory region and proposals of how heart muscle contraction may be affected as a result of these modifications. The discussion within this chapter will not include references to the solved crystal structure of the 'core' troponin complex (1J1D.pdb & 1J1E.pdb) (*28*), as this will be further addressed in later chapters of this thesis (Chapters VI, VII and VIII).

EXPERIMENTAL PROCEDURES

IV-A. Preparation of cCTnC protein and synthetic peptides

The engineering of the expression vector for the cCTnC (90-161) and expression/purification of the recombinant protein in *E. coli* BL21 DE3 (pLysS) cells (^{15}N) are as previously described in Chapter III of this thesis (*27, 29, 30*). Synthesis of peptides cIp, cIp-R145, cIp-R144G, cIp-T142Phos and cIp-R144G-T142Phos, cIp-K139G, cIp-R140G and cIp-G4 (K139G, R140G, R144G, R145G) are as previously described as for wild-type cIp in Chapter III of this thesis (*27, 31*). The design of all mutant synthetic peptides of wild-type cIp were synthesized via standard peptide synthetic procedures for titration studies against cCTnC: cIp-K139G; acetyl-**TQKIFDLRGKFGRPTLRRVR**-amide, cIp-R140G; acetyl-**TQKIFDLRGKFKGPTLRRVR**-amide, cIp-T142Phos; acetyl-**TQKIFDLRGKFKRPT$_p$LRRVR**-amide, cIp-R144G-T142Phos; acetyl-**TQKIFDLRGKFKRPT$_p$LGRVR**-amide, cIp-R144G; acetyl-**TQKIFDLRGKFKRPTLGRVR**-amide, cIp-R145G; acetyl-**TQKIFDLRGKFKRPTLRGVR**-amide, cIp-G4 (K139G, R140G, R144G, R145G); acetyl-**TQKIFDLRGKFGGPTLGGVR**-amide. Following synthesis and purification, all peptides were verified by MALDI-TOF mass spectrometry and lyophilized twice to remove all traces of organic solvents.

IV-B. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-K139G

5.44mg of ^{15}N-cCTnC was dissolved in 575*u*L of NMR buffer (100mM KCl, 10mM Imidazole, pH=6.7, 0.001% NaN$_3$, 17mM CaCl$_2$), sterile filtered and 500uL was aliquoted into a clean 5mm NMR tube. 7.66mg of solid, lyophilized cIp-K139G peptide was dissolved in 60*u*L of NMR buffer and centrifuged in a 1.5mL sterile, 0.22*u*m Nylon spin-filter (Costar, Corning Incorporated) at 3K rpm for 10 minutes to remove all impurities and any undissolved sample. The titration was composed of a total of 17 data points in which there was 5 separate additions of 1*u*L of the clp-K139G solution added directly to the NMR tube, followed by 3 additions of 1.5*u*L, 2 additions of 2*u*L, 2 additions of 3*u*L, and single additions of 4, 5, 8, 10 and 15*u*L respectively. After every titration point, 1*u*L of the resulting titrated solution was removed from the NMR tube and used for amino acid analysis. There was a small alkaline shift in pH during the titration, which was corrected for with small amounts of 0.1M HCl. The change in cCTnC•2Ca^{2+} concentrations due to aliquot additions during the titration was taken into account during data analysis. Both 1D ^{1}H-NMR and 2D ^{1}H/^{15}N-HSQC spectra were acquired at every titration point.

VI-C. Titration of ^{15}N-cCTnC•2Ca^{2+} with clp-R140G

7.45mg of ^{15}N-cCTnC and 9.57mg of clp-R140G were separated weighed out and prepared in an identical manner for that described for the clp-K139G titration, except that the clp-R140G peptide was dissolved in 42*u*L of NMR buffer. The titration was composed of a total of 16 data points in which there was 6 separate additions of 1*u*L of the clp-R140G solution added directly to the NMR tube, followed by 3 additions of 2*u*L, 2 additions of 3*u*L, and single additions of 5, 6 and 12*u*L respectively. After every titration point, 1*u*L of the resulting titrated solution was removed from the NMR tube and used for amino acid analysis. There was a similar alkaline shift in pH during the titration as observed for clp-K139G, which was again corrected for with small amounts of 0.1M HCl. The change in cCTnC•2Ca^{2+} concentrations due to aliquot additions during the titration was taken into account during data analysis. Both 1D ^{1}H-NMR and 2D ^{1}H/^{15}N-HSQC spectra were acquired at every titration point. During amino acid analysis, glass tubes containing titration points 1, 2, 5 and 14 ruptured during heated acid hydrolysis and the samples were unrecoverable, therefore for data analysis only 12 titration points were used for calculating thermodynamic values.

VI-D. Titration of ^{15}N-cCTnC•2Ca^{2+} with clp-R144G

To a NMR tube containing a 500*u*L sample of 0.92mM ^{15}N-cCTnC•2Ca^{2+} (prepared in an identical manner to that of clp-K139G) were added aliquots of 1*u*L of a 41mM clp-R144G stock solution for the first 10 titration points. Aliquots of 2*u*L were added for the next 4 titration points, followed by consecutive addition of 4*u*L, 10*u*L, 25*u*L, and 40*u*L for the last four titration points. The sample was thoroughly mixed after each addition. After every titration point, 1*u*L of the resulting titrated solution was removed and used for amino acid analysis. The change in cCTnC•2Ca^{2+} concentration due to increased volume during the titration was taken into account for data analysis, the change in pH for clp-R144G addition was negligible. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC spectra were acquired at every titration point.

VI-E. Titration of ^{15}N-cCTnC•2Ca^{2+} with clp-R145G

To an NMR tube containing a 500*u*L sample of 0.29mM ^{15}N-cCTnC (prepared in an identical manner to that of clp-K139G) were added aliquots of 1*u*L of 14.9mM clp-R145G stock solution for the first 9 titration points. Aliquots of 2*u*L were added for the next 2 titration points, followed by additions of 3*u*L, 5*u*L, and 7*u*L to ensure complete clp-R145G saturation. The sample was thoroughly mixed after each addition. After every titration point, 1*u*L of the resulting titrated solution was removed and used for amino acid analysis. The change in cCTnC•2Ca^{2+} concentration due to increased volume during the titration was taken into account for data analysis, the change in pH for clp-R145G addition was negligible. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC spectra were acquired at every titration point.

VI-F. Titration of ^{15}N-cCTnC•2Ca^{2+} with clp-T142Phos

To an NMR tube containing a 500*u*L sample of 0.40mM ^{15}N-cCTnC (prepared in an identical manner to that of clp-K139G) were added aliquots of 1*u*L of 16.6mM clp-T142Phos stock solution for the first 12 titration points. Aliquots of 2*u*L were added for the next 2 titration points, followed by aliquots of 3*u*L for 3 titration points, and two additions of 4*u*L to ensure complete cIp-T142Phos saturation. The sample was thoroughly mixed after each addition. After every titration point, 1*u*L of the resulting titrated solution was removed and used for amino acid analysis. The change in cCTnC•2Ca^{2+} concentration due to increased volume during the titration was taken into account for data analysis, the change in pH for cIp-T142Phos addition was negligible. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC spectra were acquired at every titration point.

VI-G. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-R144G-T142Phos

To a NMR tube containing a 500*u*L sample of 0.49mM ^{15}N-cCTnC•2Ca^{2+} (prepared in an identical manner to that of cIp-K139G) were added aliquots of 1*u*L of a 49mM cIp-R144G-T142Phos stock solution for the first 6 titration points. Aliquots of 1.5*u*L, 1.5*u*L, 2*u*L, and 2*u*L were added consecutively for the next 4 titration points, followed by consecutive addition of 3*u*L, 4*u*L, 10*u*L, and 10*u*L for the last four titration points. The sample was thoroughly mixed after each addition. After every titration point, 1*u*L of the resulting titrated solution was removed and used for amino acid analysis. The change in cCTnC•2Ca^{2+} concentration due to increased volume during the titration was taken into account for data analysis, the change in pH for cIp-R144G-T142Phos addition was negligible. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC spectra were acquired at every titration point.

VI-H. Titration of ^{15}N-cCTnC•2Ca^{2+} with cIp-G4 (K139G, R140G, R144G, R145G)

6.75mg of ^{15}N-cCTnC and 7.35mg of cIp-G4 were each weighed out and prepared in an identical manner for that described for the cIp-K139G titration, except that the cIp-G4 peptide was dissolved in 80*u*L of NMR buffer. The titration was composed of a total of 14 data points in which there was 4 separate additions of 1*u*L of the cIp-G4 solution added directly to the NMR tube, followed by 2 additions of 2*u*L, and single additions of 3, 4, 5, 7, 10, 15 and 25*u*L respectively. After every titration point, 1*u*L of the resulting titrated solution was removed from the NMR tube and used for amino acid analysis.

There was a similar alkaline shift in pH during the titration as observed for clp-K139G, which was again corrected for with small amounts of 0.1M HCl. The change in cCTnC•2Ca^{2+} concentrations due to aliquot additions during the titration was taken into account during data analysis. Both 1D ^{1}H-NMR and 2D ^{1}H/^{15}N-HSQC spectra were acquired at every titration point.

VI-I. NMR Spectroscopy

All NMR spectral data were obtained using a Unity INOVA 500MHz spectrometer at 30°C. All 1D ^{1}H-NMR spectra were acquired using a sweep width of 6000Hz with 32 transients. All 2D {^{1}H, ^{15}N}-HSQC spectra were acquired using the sensitivity-enhanced gradient pulse scheme developed by Lewis E. Kay and co-workers (*32*, *33*). The ^{1}H and ^{15}N sweep widths were 6000Hz or 7000Hz and 1500Hz, with 16 transients and 128 increments (1024 points in the directly detected dimension), respectively. Spectral processing of all {^{1}H, ^{15}N}-HSQC was performed using the software package NMRPipe by Frank Delaglio and co-workers (*34*) using zero-filling and backward linear prediction on all data sets, and spectral viewing, assignments using the software package nmrView by Bruce Johnson and co-worker (*35*). All samples were referenced directly (indirectly for ^{15}N dimension) to a 0.2mM internal standard of 2,2-dimethyl-2-silapentane-5-sulfonate (Cambridge Isotope Laboratories).

RESULTS

IV-A. Titration of cIp-G4 peptide with the C-lobe of cTnC

A small amount of white precipitate was observed for the titration of cIp-G4 with the C-lobe of cTnC after the first addition, and level of precipitation increased as increasing levels of cIp-G4 were added. The sample was centrifuged for 10 minutes at 8K rpm to remove all precipitate from solution prior to taking a 1*u*L sample for each amino acid analysis, however there was an inherently large error associated with this titration.

IV-B. The consequences of cIp mutation on cCTnC binding

Peptide Name	Sequence *
cIp	TQKIFDLRGKFKRPT LRRVR
cIp-K139G	TQKIFDLRGKFGRPT LRRVR
cIp-R140G	TQKIFDLRGKFKGPT LRRVR
cIp-T142Phos	TQKIFDLRGKFKRPT$_p$LRRVR
cIp-R144G	TQKIFDLRGKFKRPT LGRVR
cIp-R144G-T142Phos	TQKIFDLRGKFKRPT$_p$LGRVR
cIp-R145G	TQKIFDLRGKFKRPT LRGVR
cIp-G4	TQKIFDLRGKFGGPT LGGVR

Figure IV-1. **Mutagenesis of cIp peptides used to monitor the electrostatic binding effects of cIp to the C-lobe of cTnC**. The variable region (Lys139 to Arg145) of cIp is indicated by an asterisk (*), with basic residues coloured blue, threonine coloured green and hydrophobic/unaltered coloured yellow. Sites of mutation(s) and phosphorylation are indicated in bold and red. The mutant peptide cIp-G4 contains four distinct mutations (K139G, R140G, R144G, R145G).

Figure IV-1 indicates the mutations performed on the inhibitory region for binding studies to the C-lobe of cTnC. To demonstrate the effects of electrostatic interactions, mutational analysis of cIp binding to the C-lobe was monitored using 2 dimensional {^{1}H, ^{15}N}-HSQC NMR spectroscopy of the backbone resonances of ^{15}N-cCTnC upon the addition of peptide. The backbone assignments of the C-lobe in the calcium saturated state have been completely assigned and previously reported (see Figure III-4, Chapter III, page 74), and were used as starting points to monitor protein-peptide chemical shift changes (*3, 27*).

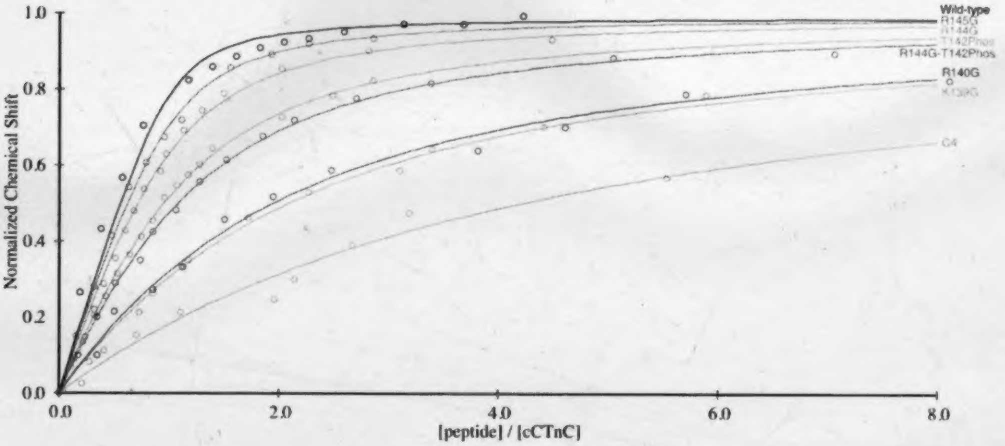


Figure IV-2. **Dissociation constants (K_D) for various cIp mutant peptides with the C-lobe of cTnC**. The eight different peptides tested for binding studies are colour indicated for clarity. A more linear best-fit line represents a lower affinity of the peptide for the C-lobe. The K_D values were generated using the software program XCRVFIT (*36*) and are reported in *Table IV-1*.

For all measured peptides, the chemical shift perturbations were within the fast exchange timescale NMR limit, so that each cross-peak corresponded to the weighted average of the bound and free chemical shifts. Linear movement of all cross-peaks indicated that only two species existed in solution. Resonances undergoing large backbone amide ^{1}HN and/or ^{15}N chemical shift perturbations were monitored, measured and normalized to 1.0 for each peptide, as previously described for the cIp wild-type peptide in Chapter III (*27*). Chemical shift data was fit to Equation [1] to obtain a dissociation constant (K_D) for the titration of each individual peptide (denoted as *). All peptide titration curves were then re-normalized relative to one another to produce the binding curve displayed in Figure IV-2, which spans from 0 to 8 molar equivalent of peptide additions to the C-lobe and chemical shift values approach 1.0 as the X-axis approaches ∞. In keeping with the methodology used in deriving the experimental dissociation constant for the wild-type peptide (see Figure III-7, Chapter III, page 78), only resonances Ile128, Thr129, Gly159 and Val160 were used for induced chemical shift analysis. The experimental dissociation values obtained from the binding curve are reported in *Table IV-1.*

$$\text{cCTnC}\bullet 2\text{Ca}^{2+} + \text{cIp}^* \leftrightarrow \text{cCTnC}\bullet 2\text{Ca}^{2+}\bullet\text{cIp}^* \quad [1]$$

Table IV-1

Dissociation constants of various cIp peptides binding to C-lobe of cTnC

Peptide Name	K_D (*u*M)	Affinity Reduction[a]
cIp – wild type	31 ± 11	1.0
cIp – R145G	123 ± 17	4.0
cIp – R144G	193 ± 21	6.2
cIp – T142Phos	451 ± 10	14.6
cIp – R144G-T142Phos	581 ± 12	18.7
cIp – R140G	1488 ± 150	46.7
cIp – K139G	1582 ± 133	51.0
cIp – G4	3667 ± 715	118.3

[a] Calculated using the K_D values of each individual mutant peptide compared to the wild-type peptide.

Overlay of the {^{1}H, ^{15}N}-HSQC titration spectra for each individual peptide reveals that although the peptides bind to the C-lobe with greatly differing dissociation constants, chemical shift changes to the C-lobe remained virtually unchanged throughout the scanning mutagenesis study, therefore suggesting that the binding event remains constant and the peptide does not possess an altered form in the mutated state. Figure IV-3 shows the induced chemical shift changes for only the residues that underwent greater than one standard deviation of the average for the wild-type cIp peptide (see Figure III-5, Chapter III, page 76). Due to the common trend of induced chemical shift changes to the C-lobe, the conformational and/or secondary structural changes to either domain is assumed to be the same for the binding of each peptide. This is a critical assumption of any scanning mutagenesis approach, such that the binding changes in free energy observed are due only to the residue(s) mutated/modified, and not any concomitant structural changes.

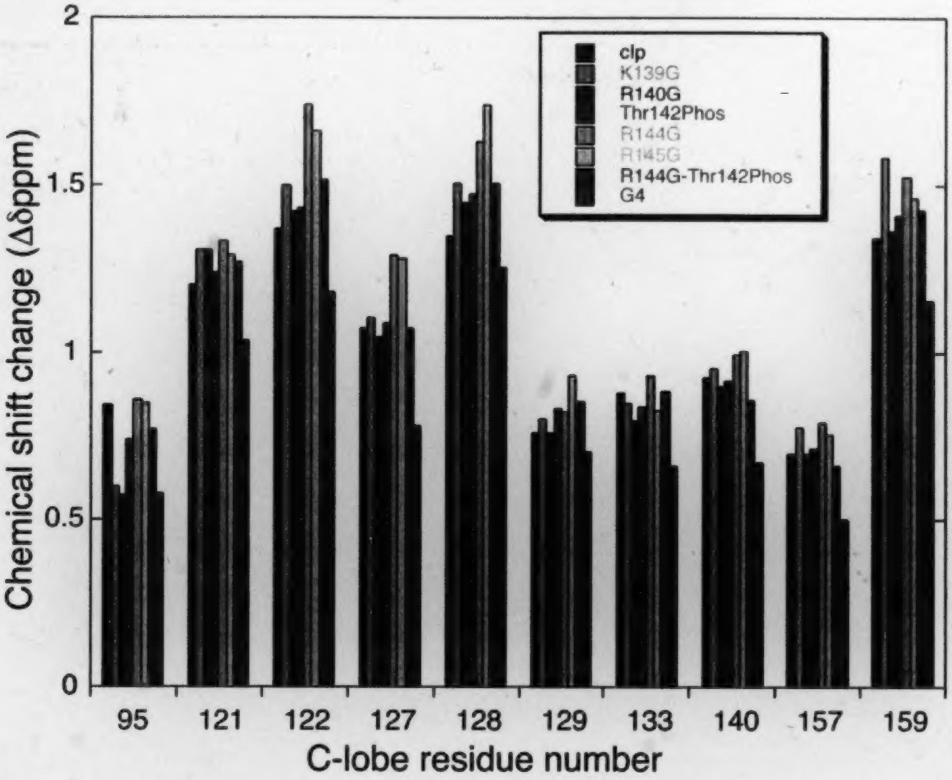


Figure IV-3. **Induced C-lobe chemical shift changes by cIp peptides**. Peptides are color indicated, and within one standard deviation of one another for each specified residue.

Based upon the measured K_D values, a trend regarding the placement of a charged residue on the inhibitory region versus the dissociation constant is observed. The affinity between the two domains is strongly charge dependant, which can be attributed to electrostatic residues making positive attractions between the two domains. As the nullification of a charged residue moves from the C-terminal tail of cIp (i.e. R145G, $K_D \sim$ 123*u*M), towards the central and N-terminal regions of cIp, which interact more closely with the C-lobe (i.e. R140G, $K_D \sim$ 1488*u*M), the binding affinity becomes weaker and weaker. The observed dependence of the glycine-scanning mutagenesis was so sensitive to the electrostatic potential(s) and distance(s) that the nullification of an arginine residue only 5 residues apart (R140G compared to R145G) is sufficient to impart a difference in binding affinity which is on the order of two orders of magnitude, when compared to the wild-type peptide. The positioning of each mutation in relation to the published structure (1OZS) of the inhibitory region bound to the C-lobe is shown in Figure IV-4A. Upon visual inspection of the structure, it becomes clear how a nullification of a basic residue to a non-charged residue may have a strong effect when considering proximity to the acidic C-lobe.

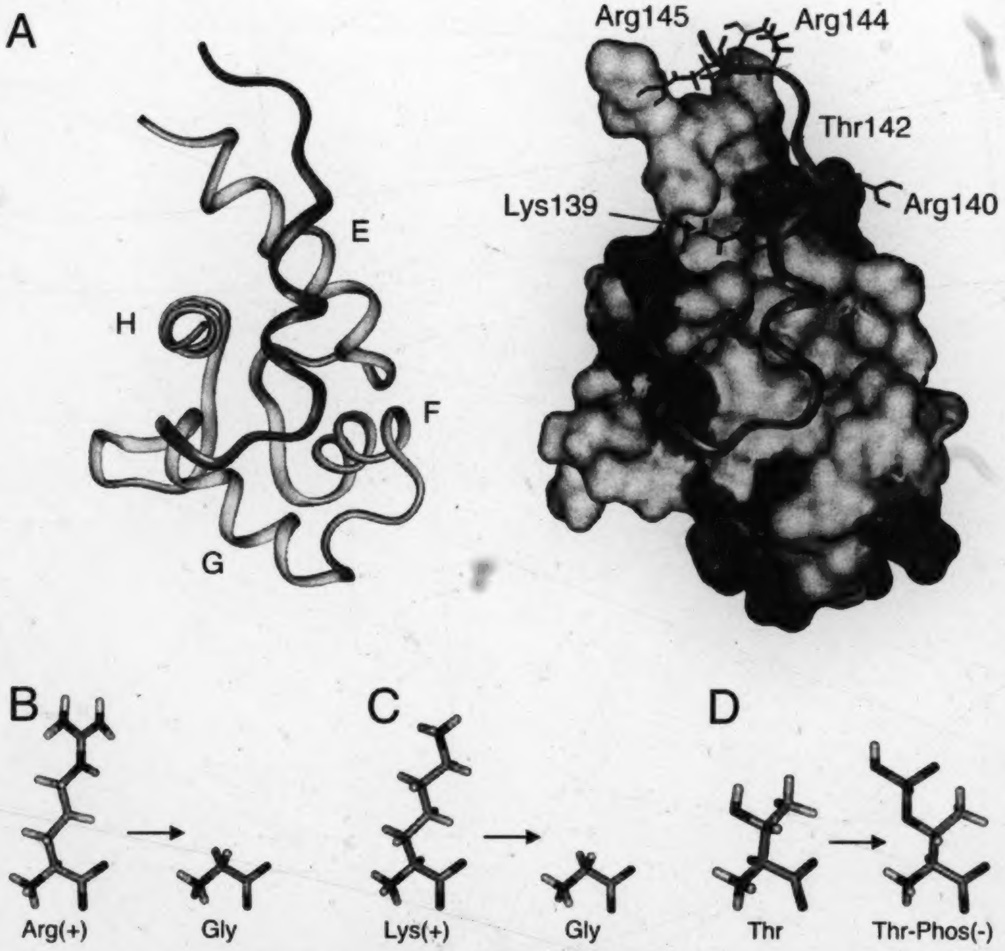


Figure IV-4. **Mutation and phosphorylation of the inhibitory region**. (A) Left - Ribbon diagram of a single model from the NMR ensemble structure 1OZS, displaying the helices of the C-lobe (yellow) and orientation of the inhibitory region (grey). Right - Connolly surface representation of the C-lobe with all acidic residues coloured in red and residues undergoing mutation/phosphorylation are indicated in blue. (B-D) Stick representation of chosen residues to indicate all mutational and phosphorylation events.

The choice to mutate all basic residues to glycine was pursued to follow the biological FHC condition on Arg144 (see Figures IV-4B to IV-4D), yet considerations other than just charge placement/elimination must be addressed. The nullification of the positive charge on the Arg/Lys side-chains that was achieved by glycine mutagenesis, resulted in the loss of the hydrophobic methylene side-chain carbon groups (i.e. $-CH_2-CH_2-CH_2-CH_2-$), which may have affected side-chain packing required for proper complex formation. However no hydrophobic inter-domain NOEs were observed from any of the mutated cIp side-chain residues to the C-lobe in the published structure. A mutation to a differing residue such as leucine would succeed in eliminating the positive charged side-chain and leave a hydrophobic side-chain that is similar in length and size to both lysine and arginine. However this was not pursued due to time considerations, cost of peptide synthesis, unwanted consequences to peptide solubility, and the possible dramatic effect regarding the nature of the binding due to the close proximity to the C-lobe hydrophobic face (see yellow region – Figure IV-4A). Glycine is the only residue of the 20 common amino acids which does not contain a side-chain, which increases the potential for flexibility upon domain association. In particular the quadruple glycine mutant cIp-G4 may possess additional flexibility, which may impact the binding constant and energetic potentials.

IV-C. Distance effects on electrostatic attractive forces

The placement of C-lobe acidic residues is prominently displayed in Figure IV-4A, with the residues of the E-helix in the closest proximity to the basic residues undergoing mutation within cIp. The exact positioning of the side-chains with relation to one another is unfortunately not completely known, as the NMR solution structure (1OZS) contained no structural restraints for residues Arg140 to Arg 147 with the C-lobe, therefore appearing as if the C-terminal tail of cIp was completely labile in solution (see Figure III-8, Chapter III, page 82). However based upon ^{15}N-NMR relaxation data of ^{15}N-cIp, we know that the inhibitory peptide binds rigidly and tightly to the C-lobe throughout the entire length of its sequence. Based upon the potentials for charged side-chains within the complex, there is high probability that electrostatic interactions between the C-terminal region of the inhibitory peptide and the C-lobe will be occurring, and were simply not resolved in the NMR experimentation. However, electrostatic forces are greatly distance dependant and extend their respective influences well beyond the ~5Å NOE limit ($1/r^6$) found in NMR experiments. Coupled with the solvent exchangeability of the protons present on Lys (NH_3+) and Arg (H_2N-C-NH_2+) side-chains, it was therefore not surprising that no NOE's were observed between the two domains during NMR structure elucidation.

Obtaining the potential energy (***E***) of each salt-bridging pair and relating it to the binding affinities is the ultimate goal of this chapter. However, based solely on the solution structure of the binary complex, it is impossible to determine or predict with high certainty which acidic residue might be pairing with a basic partner. To quantitate the electrostatic interactions occurring within the binary complex, the distances involved between the two domains must be determined. This presents a large problem due to the large degree of 'flexibility' within the reported NMR binary solution structure. Therefore a number of assumptions have to be made. To obtain reliable distances of the electrostatic entities, it was decided that the "average" distance between charged groups within the reported family of 30 NMR ensemble structures (1OZS) might be a close estimate of the actual distances occurring. In keeping with the knowledge that a charged particle in solution can exert an electric potential across relatively large distances (1/r, ≥20Å), it was decided to calculate the potential energy from each point mutation on the cIp peptide to that of all acidic residues within the C-lobe, as most C-lobe residues are closer than 30Å to any given charged residue on the inhibitory peptide. Consideration must be taken towards the interactions of both cIp acidic and basic residues (total of 9 throughout cIp, plus one phosphorylated Thr residue) with both acidic and basic C-lobe residues (total of 32 in the C-lobe), as well as interactions of additional non-mutated residues within cIp interacting independently with the C-lobe. Therefore, calculations included both attractive and repulsive forces occurring between both C-lobe and the cIp residues. The calculations of potential energy (Equation [2]) follow the mutational analysis shown in Figure IV-5, where h is the number of structures in the pdb ensemble, i is a point charge on cIp and j is a point charge on the C-lobe. Phosphate groups were considered to have an elementary charge of –1 in the buffered solution at pH 6.7, based on the pka's of phosphoric acid, which are 2.15, 7.20 and 12.15 respectively.

Region of clp peptide considered	Wt	K139G	R140G	T142Phos	R144G	R145G	R144G-T142Phos	G4
Arg145	+	+	+	+	+	**0**	+	**0**
Arg144	+	+	+	+	**0**	+	**0**	**0**
Leu143	0	0	0	0	0	0	0	0
Thr142	0	0	0	**-**	0	0	**-**	0
Pro141	0	0	0	0	0	0	0	0
Arg140	+	+	**0**	+	+	+	+	**0**
Lys139	+	**0**	+	+	+	+	+	**0**

Figure IV-5. **Mutational analysis of the inhibitory peptide.** Schematic diagram of placement of charges on clp which were used in calculating the potential energy of interaction with the C-lobe. A (+) symbol indicates an elementary particle point charge of 1.602E-19 C, a (-) symbol indicates a point charge of –1.602E-19 and a **0** indicates a zero net point charge.

$$\boldsymbol{E} = \frac{1}{30}\sum_{h=1}^{30}\sum_{i=1}^{10}\sum_{j=1}^{32}\frac{Q_iQ_j}{4\pi\varepsilon_0\varepsilon r_{ij}} \qquad [2]$$

For free energy calculations (ΔG), a temperature value (T) of 303.15°K (30°C) and a gas constant (R) of 8.315 $JK^{-1}mol^{-1}$ were assumed (Equation [3]), with K_D values obtained from analysis of HSQC spectra using Equation [1]. For potential energy calculation, an assumed dielectric constant (ε) of 80 for the solvent system was chosen in addition to the permittivity of a vacuum (ε_0) constant of 8.854×10^{-12} $C^2J^{-1}m^{-1}$ and an elementary particle charge (Q) value of $\pm 1.602 \times 10^{-19}$ C. All side chain moieties were treated as assumed point-charges, with distances derived from the carboxylate carbons for Asp/Glu residues (IUPAC nomenclature - atoms CG/CD), the amine nitrogen for Lys residues (NZ), the carbon of the guanidinyl group for the Arg residues (CZ), and the side-chain hydroxyl oxygen for Thr residues (OG1). The thermodynamic values for free energy of binding and potential energy are detailed in *Table IV-II*. The potential energy values were calculated in Equation [2] and the free energy values (ΔG) were derived from the binding constants in *Table IV-I* and Equation [3].

$$\Delta G = -RT \ln K_{\text{association}} = RT \ln K_{\text{Dissociation}} \quad [3]$$

Table IV-II

Thermodynamic data of clp peptides binding to C-lobe of cTnC

Peptide Name	Potential Energy (J/mol)[a]	$\Delta\Delta E$ (J/mol)[b]	Free Energy (J/mol)	$\Delta\Delta G$(J/mol)[b]
clp – wild type	-73830 ± 5461	0	-26169	0
clp – R145G	-66362 ± 5028	7468	-22695	3474
clp – R144G	-65158 ± 4659	8672	-21695	4474
clp – T142-phos	-63592 ± 5172	10238	-19420	6749
clp – R144G/T142-phos	-54921 ± 4303	18909	-18781	7388
clp – R140G	-61665 ± 4772	12165	-16479	9690
clp – K139G	-61846 ± 5180	11984	-16256	9913
clp – G4	-33482 ± 3613	40348	-14137	12032

[a] Derived from the average distance is taken from the 30 lowest energy calculated ensemble of the solved NMR structure 1OZS (*16*), with all charged groups treated as point charges. The error is defined as one standard deviation of the mean for all calculations.

[b] Values obtained for $\Delta\Delta E$ and $\Delta\Delta G$ were calculated from each peptide relative to the wild-type.

The free energy component (ΔG) for any reaction can be broken up into the standard free energy equation, which includes enthalpic (ΔH) and entropic (ΔS) contributions, which is shown in Equation [4].

$$\Delta G = \Delta H - T\Delta S \quad [4]$$

The enthalpic contribution to the free energy can be considered the summation of a potential energy term ($\Delta \mathbf{\mathit{E}}$), plus a work term, which is shown in Equation [5].

$$\Delta H = \Delta \mathbf{\mathit{E}} + \text{work} \quad [5]$$

When equation [5] is placed back into equation [4], an expanded formula is derived for ΔG. This formula states that ΔG is simply equal to ΔE plus additional values, which hold constant (C) for a given reaction, which are shown in Equations [6] and [7]. Considering that the potential energy (ΔE) is a product of $1/r_{ij}$, as defined earlier in Equation [2], a linear relationship should be observed for a plot of free energy (ΔG) versus potential energy (ΔE), which should reflect a positive slope of $(Q_iQ_j / 4\pi\varepsilon\varepsilon_0)$.

$$\Delta G = \Delta E + \text{work} - T\Delta S \qquad [6]$$

$$\Delta G = \left(\frac{Q_iQ_j}{4\pi\varepsilon_0\varepsilon}\right)\frac{1}{r_{ij}} + C \qquad [7]$$

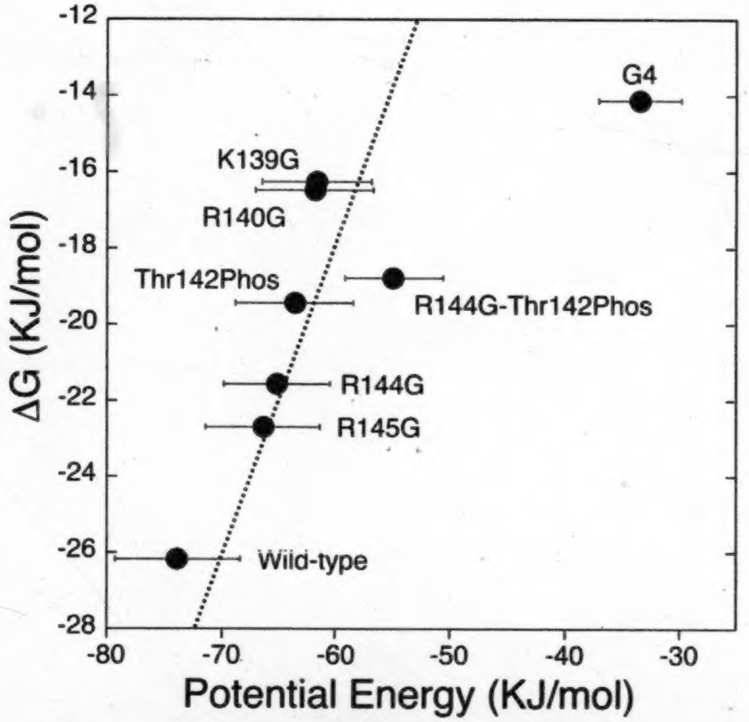


Figure IV-6. **Linear relationship of free energy of binding (ΔG) versus potential energy (ΔE).** The various clp peptides are as indicated in the plot, with error bars for potential energy displayed in red. All potential energy values are negative which implies an attractive electrostatic force between the domains.

The data reported for all performed cIp mutations and/or phosphorylation events in *Table IV-II* was used in the construction of Figure IV-6. The plot reveals a near linear relationship for the binding of the modified cIp peptides with the C-lobe. Two of the tested peptides were outside of the error limits for the line of best fit, namely R144G-Thr142Phos and G4. The large deviation observed for peptide G4 is not surprising, considering the binding constant (K_D = 3667*u*M) was higher than the concentration of the NMR sample tested in the titration, thereby yielding a titration with a high intrinsic associated error. In addition, the mutation of four basic residues to Gly may have slightly affected the side-chain placements of the other basic residues on cIp, thereby adding additional error in the calculation of distances between perspective point charges for elucidation of potential energy. In Figure IV-6, the line of best fit does not include the value for the G4 peptide. The measurements for modified cIp peptides strongly support the earlier hypothesis in Chapter III that the reported binary complex is indeed highly dependant on electrostatic interactions for association, and gives credibility to the published structure.

DISCUSSION

Based upon the NMR solution structure presented in Chapter III, the binding of the inhibitory region to the C-lobe is mediated via a limited number of hydrophobic contacts and a majority of undefined electrostatic interactions (*16*). In this study, I have directly addressed potential electrostatic interactions using designed, modified versions of the wild-type inhibitory peptide. All the studied mutations involved a change to a more negatively charged cIp region, which was deemed responsible for electrostatic repulsion (or lack of attraction) between the cIp region and the acidic C-lobe (pI=4.0), which altered binding affinities to lower values. The binding affinities for the peptides (*Table IV-I*) behaved in a manner as previously predicted in Chapter III, such that as the mutation gets closer in space to the C-lobe, the binding affinity becomes weaker, so much so that the mutant cIp-K139G binds at almost two orders of magnitude weaker than that reported for cIp-R145G (*27*). The observed experimental results correlate nicely with the well-known consequence of varying the distance between electrostatic attracting (or repulsing) point charges in an aqueous environment (see Equation [2]).

Although all of the designed peptides of the inhibitory region attempted to probe electrostatic environments and the resulting consequence of the mutation, two of the tested peptides (Thr142Phos and R144G) are known to occur *in vivo* that can alter muscular contraction within the sarcomere. They are therefore directly relevant to bio-medical research in attempting to correlate *in vitro* results with *in vivo* observations. The mutation cIp-R144G is of medical importance as it is a known cause of Familial Hypertrophic Cardiomyopathy (FHC), a disease within humans causing a thickening of the left heart ventricle resulting in pre-mature death due to myocardial infarction (*18, 19*). Modification of Thr142 via phosphorylation by PKÇ can primarily act as a downward modulation of cardiac contractility by decreasing actomyosin ATPase activity.

The inhibitory region of cTnI is prone to missense mutations in diseased heart. Mutations such as R144G or R144Q that co-segregate with FHC are located in this highly conserved region (*19*). The R144G mutation constitutes a change in charge, from a positive to a neutral entity. The functional abnormalities revealed a reduced ability of R144G mutants to inhibit actomyosin ATPase and an increase in Ca^{2+} sensitivity of actomyosin ATPase regulation (*37-41*). If these functional differences manifest themselves *in vivo* there will be a resulting impairment of relaxation of cardiac muscle, and this altered contractility may provide a hypertrophic stimulus eventually leading to cardiac dysfunction. This is supported by a mouse model study that shows that at the whole organ level, contractile function is enhanced but relaxation is compromised in mouse harboring the R144G mutation (*38*). Because cTnI toggles between cTnC and actin-tropomyosin, it seems likely that an important mechanism of cTnI-induced cardiac dysfunction by the R144G mutation lies in altered cTnI interactions with cTnC, actin-tropomyosin or both. The binding of cTnI to actin-tropomyosin inhibits actomyosin ATPase activity, depresses cross-bridge cycling, and prevents contraction. The inhibition is released as cTnI binding shifts from actin-tropomyosin to cTnC, an action favored by an increase in cytosolic Ca^{2+} concentration during systole, when Ca^{2+} binds to the N-domain of cTnC. Thus, it seems that the R144G mutant within cTnI either has reduced affinity for actin-tropomyosin or enhanced affinity for cTnC. The net effect would be the diminished inhibition of actomyosin ATPase activity. Interestingly, my results indicate that the affinity of cIp for cCTnC•2Ca^{2+} is reduced by ~6-fold by the mutation R144G. In light of the above argument, this is somewhat conflicting in that the affinity of cIp for troponin C is actually decreased instead of increased, as would be predicted from the *in vivo* observations. However based solely on an electrostatic viewpoint, a reduction in affinity would be predicted.

The answer to this confliction seems to stem from three possible explanations. The first is that this study was performed by NMR spectrometry, a biophysical experimentation tool where the molecular weight of the interacting domains is intentionally kept to a minimum for optimal line-shape and assignments. The fact that I only used a small interacting portion of the troponin complex instead of the entire troponin molecule, or even whole muscle fiber is a limitation in my experimentation. Can my results from this study explain all observed *in vivo* observations? The answer is simply *no*. Optimal studies would include an ability to monitor at atomic resolution the entire myofibril in an altered-state of R144G compared to the wild-type, however the ability to perform these advanced studies are not yet practical with the current technology available and we are thus dependant on piecing together data from smaller pieces such as my study to elucidate the functioning of the larger complex. A second possible explanation to the apparent confliction is that the majority of the consequences of the R144G mutation within the myofibril are directed to the inhibition of actomyosin ATPase activity. This result seems more likely as the reported observations simply do not explain the *in vivo* observations. However since there is no high-resolution data available regarding the interactions of cTnIs inhibition of actomyosin ATPase activity, we are currently limited to predictions of the nature of the reduced inhibition. The third possible explanation to the conflict builds further on the first explanation, which is simply that the NMR structure that we based these measurements on is in some form altered in the full-length complex. This is a very real possibility that will be further explored in Chapters VI, VII and VIII of this thesis.

This chapter strictly focuses on only a Thr142 phosphorylation event within the inhibitory region. Chapter VI of this thesis will address phosphorylation and regulation by Ser41/Ser43. The relative contribution of the phosphorylation of these three sites to the accelerated relaxation and decreased actomyosin ATPase activity has been the focus of several recent studies. For example, Burkart and co-workers reported that phosphorylation of Ser41/Ser43 dominates regulation of the level of maximum tension, while phosphorylation of Thr142 appears to be required for regulation of thin filament sliding speed (*42*). The data presented here demonstrated that it is the phosphorylation of Thr142 that plays a major role in destabilizing the cTnI-cTnC interaction, which may lead to diminished actomyosin ATPase activity and depressed contractility. The binding affinity of Thr142Phos was shown to have a 14.6-fold reduction in the binding affinity of the inhibitory region with the C-lobe. In terms of electrostatics, the data reported for the phosphorylation of Thr142 better fits the observed *in vivo* studies, however this reported value may still suffer, in part, from the three conflicting answers proposed for the R144G mutation. Regardless of the mentioned limitations, the understanding of how phosphorylation plays a key role in muscular contraction is immediately medically relevant as it affects everyone within the world, versus the R144G mutation that only affects a small portion of the population. Nonetheless, the dramatic affinity reduction of cTnI for cTnC by the phosphorylation of T142 strongly suggests that kinases such as PKC can modulate cardiac muscle contraction by modifying myofilament protein-protein interactions, contributing to depressed contractility in certain cases and compromised relaxation in others.

The focus of this study was to probe the electrostatic interactions occurring between the inhibitory region and the C-lobe, and correlate charge placement on the inhibitory region with affinity reduction arising from mutagenesis/phosphorylation experimentation, as per the solved NMR solution structure presented in Chapter III of this thesis. When the structure of the binary complex was published, I was somewhat concerned as it did not closely match the solved crystal structure of the troponin complex (1J1E.pdb). Naturally, I was concerned that a mistake may have been made during NMR structural assignments. However, further examination of the data that generated the published structure yielded the same conclusion that the presented binary complex was correct. Analysis of the electrostatic environment that I performed in this chapter supports the validity of the NMR solution structure of the binary complex. It should be noted however that the conformation of the inhibitory peptide may in part be compromised in the full-length troponin complex, a theory that will be examined and discussed in Chapters VI, VIII and VIII in this thesis.

REFERENCES

1. Geeves, M. A., and Holmes, K. C. (1999) *Annu Rev Biochem 68*, 687-728.
2. Gordon, A. M., Homsher, E., and Regnier, M. (2000) *Physiological Reviews 80*, 853-924.
3. Sia, S. K., Li, M. X., Spyracopoulos, L., Gagné, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) *J. Biol. Chem. 272*, 18216-18221.
4. Spyracopoulos, L., Li, M. X., Sia, S. K., Gagné, S. M., Chandra, M., Solaro, R. J., and Sykes, B. D. (1997) *Biochemistry 36*, 12138-12146.
5. Solaro, R. J., and Rarick, H. M. (1998) *Circ Res 83*, 471-80.
6. Rarick, H. M., Tu, X. H., Solaro, R. J., and Martin, A. F. (1997) *J Biol Chem 272*, 26887-92.
7. Ramos, C. H. (1999) *J Biol Chem 274*, 18189-95.
8. Campbell, A. P., and Sykes, B. D. (1991) *J Mol Biol 222*, 405-21.
9. Krudy, G. A., Kleerekoper, Q., Guo, X., Howarth, J. W., Solaro, R. J., and Rosevear, P. R. (1994) *J. Biol. Chem. 269*, 23731-23735.
10. Howarth, J. W., Krudy, G. A., Lin, X., Putkey, J. A., and Rosevear, P. R. (1995) *Protein Sci. 4*, 671-680.
11. Kleerekoper, Q., Howarth, J. W., Guo, X., Solaro, R. J., and Rosevear, P. R. (1995) *Biochemistry 34*, 13343-52.
12. Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) *FEBS Lett 453*, 107-12.
13. Abbott, M. B., Dvoretsky, A., Gaponenko, V., and Rosevear, P. R. (2000) *FEBS Lett 469*, 168-72.
14. Abbott, M. B., Gaponenko, V., Abusamhadneh, E., Finley, N., Li, G., Dvoretsky, A., Rance, M., Solaro, R. J., and Rosevear, P. R. (2000) *J Biol Chem 275*, 20610-7.
15. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.
16. Lindhout, D. A., and Sykes, B. D. (2003) *J Biol Chem 278*, 27024-34.
17. Gaponenko, V., Abusamhadneh, E., Abbott, M. B., Finley, N., Gasmi-Seabrook, G., Solaro, R. J., Rance, M., and Rosevear, P. R. (1999) *J Biol Chem 274*, 16681-4.
18. Redwood, C. S., Moolman-Smook, J. C., and Watkins, H. (1999) *Cardiovasc Res 44*, 20-36.
19. Hernandez, O. M., Housmans, P. R., and Potter, J. D. (2001) *J Appl Physiol 90*, 1125-36.
20. McDonough, J. L., Arrell, D. K., and Van Eyk, J. E. (1999) *Circ Res 84*, 9-20.
21. Murphy, A. M., Kogler, H., Georgakopoulos, D., McDonough, J. L., Kass, D. A., Van Eyk, J. E., and Marban, E. (2000) *Science 287*, 488-91.
22. D'Costa, M., Fleming, E., and Patterson, M. C. (1997) *Am J Clin Pathol 108*, 550-5.
23. Van Eyk, J. E., Powers, F., Law, W., Larue, C., Hodges, R. S., and Solaro, R. J. (1998) *Circ Res 82*, 261-71.

24. McDonough, J. L., Labugger, R., Pickett, W., Tse, M. Y., MacKenzie, S., Pang, S. C., Atar, D., Ropchan, G., and Van Eyk, J. E. (2001) *Circulation 103*, 58-64.
25. Lee, T. H., and Goldman, L. (2000) *N Engl J Med 342*, 1187-95.
26. Li, M. X., Spyracopoulos, L., Beier, N., Putkey, J. A., and Sykes, B. D. (2000) *Biochemistry 39*, 8782-8790.
27. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.
28. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
29. Chandra, M., Dong, W. J., Pan, B. S., Cheung, H. C., and Solaro, R. J. (1997) *Biochemistry 36*, 13305-13311.
30. Li, M. X., Gagné, S. M., Tsuda, S., Kay, C. M., Smillie, L. B., and Sykes, B. D. (1995) *Biochemistry 34*, 8330-8340.
31. Li, M. X., Wang, X., Lindhout, D. A., Buscemi, N., Van Eyk, J. E., and Sykes, B. D. (2003) *Biochemistry 42*, 14460-8.
32. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.
33. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.
34. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
35. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603-614.
36. Boyko, R., and Sykes, B. D. (2003) *XCRVFIT, Protein Engineering Networks of Centres of Excellence v.2.1.1*, http://www.pence.ca/software/xcrvfit/index.html.
37. Elliott, K., Watkins, H., and Redwood, C. S. (2000) *J Biol Chem 275*, 22069-74.
38. James, J., Zhang, Y., Osinska, H., Sanbe, A., Klevitsky, R., Hewett, T. E., and Robbins, J. (2000) *Circ Res 87*, 805-11.
39. Takahashi-Yanaga, F., Morimoto, S., Harada, K., Minakami, R., Shiraishi, F., Ohta, M., Lu, Q. W., Sasaguri, T., and Ohtsuki, I. (2001) *J Mol Cell Cardiol 33*, 2095-107.
40. Deng, Y., Schmidtmann, A., Redlich, A., Westerdorf, B., Jaquet, K., and Thieleczek, R. (2001) *Biochemistry 40*, 14593-602.
41. Lang, R., Gomes, A. V., Zhao, J., Housmans, P. R., Miller, T., and Potter, J. D. (2002) *J Biol Chem 277*, 11670-8.
42. Burkart, E. M., Sumandea, M. P., Kobayashi, T., Nili, M., Martin, A. F., Homsher, E., and Solaro, R. J. (2003) *J Biol Chem 278*, 11265-72.

CHAPTER V

Interactions of cRP40 and the inhibitory region of troponin I with troponin C

OVERVIEW

This chapter continues my investigation of the interactions of the C-lobe of cardiac troponin C with troponin I. In the first half of this chapter, interactions of the inhibitory region and the cRP40 region with the C-lobe will be explored using ^{15}N-NMR relaxation measurements to establish a binding preference of the two domains for a common binding face on the C-lobe (currently unpublished data). The second half of this chapter will deal exclusively with phosphorylation events within cRP40 region (Ser41/Ser43), and the effects on cardiac physiology and function. All work involving the inhibitory region was performed solely by myself, while NMR studies involving the phosphorylation effects of cRP40 was performed by Monica X. Li, with additional contributions from myself, Xu Wang, Nina Buscemi, Jennifer E. Van Eyk, and Brian D. Sykes. All phosphorylation data presented in this chapter was published in the journal *Biochemistry* on December 16th, 2003 [*Biochemistry*. 2003 Dec 16;**42(49)**:14460-8.].

The literature regarding the placement and binding preferences for both the inhibitory region and the cRP40 region is inconsistent, and as such an agreement regarding the placement, physiological role and interactions of the two domains with troponin C has not been established. My NMR data (Chapters III & IV) did not match the cardiac crystal structure of the 'core' troponin complex (1J1E.pdb) in the placement of these two domains relative to troponin C. This discrepancy has been addressed and yields the conclusion that both troponin I domains can simultaneously bind to the C-lobe. These results directly agree with my proposed model of the core troponin complex discussed in Chapter III, Figure III-18B, page 101, and a recent crystal structure of the *skeletal* troponin isoform (*1*). Additionally, my earlier studies regarding the phosphorylation of inhibitory region in Chapter IV (Thr142) is expanded with the investigation of the Protein Kinase C dependant phosphorylation of Ser41 and Ser43 within the cRP40 region.

INTRODUCTION

The cardiac troponin complex is a large, yet flexible group of individual domains that work together in the movements required for muscle contraction and relaxation. The recent wealth of structural information from various biophysical techniques in exploring the troponin complex has been a key asset in elucidating an understanding of the calcium-signaling pathway for muscle contraction via the thin filament. The structure of cardiac TnC has been solved by both crystallography and NMR spectroscopy techniques, and is the best characterized subunit of the troponin complex (*2-10*). Prior to the year 2003, structural studies of cTnI and cTnT had proven to be more elusive due to solubility and flexibility issues. The recent crystal structure of the 52kDa 'core' domain (1J1E.pdb), solved by Takeda (*3*), revealed domain positioning and orientations of cTnI and cTnT with respect to cTnC. This was a necessary evolution regarding structural studies of the thin filament contractile regulatory proteins, as past experimentation was limited to smaller fragments and proteolysis (*6, 11-24*). This 'core' domain structure has been insightful in understanding at atomic resolution the positioning of the domains with respect to one another: the coiled-coil interaction between cTnI and cTnT to form the IT arm domain as well as the positioning of the C-lobe of cTnC within the core troponin complex (*3*). However, the crystal structure lacked electron density of the critical inhibitory region of TnI ($cTnI_{137-146}$), indicating an extended linker region between the coiled-coil IT arm ($cTnI_{90-136}$) and the switch region ($cTnI_{147-163}$). The conformation and positioning of the inhibitory region in relationship to the troponin complex has long been an area of discussion and conflict, with many groups attempting to gain insights into the nature of the structural role of the inhibitory region in muscle homeostasis (*9, 19, 25-32*).

In Chapter III, I reported the NMR solution structure of the binary complex of the C-lobe of cTnC bound to cIp in the calcium-saturated state (*9*). Confirmed by ^{15}N-NMR relaxation studies, this structure revealed that the inhibitory region binds rigidly to cCTnC along the entire length of cIp, making numerous hydrophobic contacts to the F-G linker region and the E and H helices of cCTnC with numerous electrostatic interactions between the two domains. However, this reported binary complex does not correlate with the crystal structure in regards to the placement of cTnI domains; there is an area of steric clash on the hydrophobic face of the C-lobe of cTnC between the inhibitory region (binary complex) and the cRP40 region ($cTnI_{33\text{-}71}$ – crystal structure). Previous suggestions that the inhibitory region may displace the cRP40 region off the C-lobe during muscle contraction is unlikely due to the high affinity (< 1*u*M) of the cRP40 region versus the lower affinity of the inhibitory region (31*u*M) (*19, 20, 33*).

To gain an understanding of the interactions of the inhibitory and cRP40 regions for troponin C, I have utilized {^1H, ^{15}N}-HSQC NMR spectroscopy to monitor the binding effects of these regions on the C-lobe to establish a binding preference. Additionally, I have incorporated ^{15}N-NMR relaxation measurements of the C-lobe in complex with either peptide alone and/or both peptides bound. The cRP40 region is capable of displacing the inhibitory region off the hydrophobic binding face of the C-lobe, however ^{15}N-NMR relaxation reveals that the inhibitory region remains bound to the C-lobe in the presence of cRP40. The model proposed in Chapter III (Figure III-18B, page 101) is validated by the presented ^{15}N-relaxation data, when combined with the structural features of the core crystal structure. Further hypotheses regarding the flexible nature of the troponin complex are discussed in terms of regulation and functionality.

The fine-control of cardiac contraction via troponin I phosphorylation has been an underlying theme of this thesis. In this chapter, I will expand upon previous phosphorylation studies of the inhibitory region (Thr142) to include the cRP40 region of cTnI. The cRP40 domain contains two additional protein kinase C (PKC) dependant serine residues (Ser41/Ser43). Phosphorylation of cTnI appears to be of particular importance in modulating myofilament activity. The effect of cTnI phosphorylation by these kinases is primarily a downward modulation of cardiac contractility, either by decreased acto-myosin ATPase activity, which is mediated by PKC, or by enhanced Ca^{2+} dissociation from cTnC and reduced myofilament Ca^{2+} sensitivity, mediated by protein kinase A (PKA).

Using {^{1}H, ^{15}N}-HSQC NMR spectroscopy, I have monitored the binding effects of cRP40 on ^{15}N-cTnC•3Ca^{2+}, and the corresponding changes resulting from phosphorylation events at residues Ser41/Ser43. Minimal disruptions in the interaction of cRP40 with cTnC•3Ca^{2+} were observed by Ser41/Ser43 phosphorylation. This data presents as a sharp contrast to the inhibitory and switch regions, where the affinity of cIp for cCTnC•2Ca^{2+} and cSp for cNTnC•Ca^{2+} was markedly reduced by phosphorylation. The results suggest that while the modifications in the cRP40 region have minimal influence, those in the key functional cTnI regions such as cIp may have a pronounced effect on the interaction of cTnI and cTnC, which may correlate with the altered myofilament function and cardiac muscle contraction under pathophysiological conditions.

EXPERIMENTAL PROCEDURES

V-A. Preparation of cCTnC and cTnC protein and synthetic peptides

The engineering of the expression vector for the cCTnC (90-161) and expression/purification of the recombinant protein in *E. coli* BL21 DE3 (pLysS) cells (^{2}H, ^{15}N, and/or ^{13}C) are as previously described in Chapter III (*33-35*). Production and purification of cTnC is identical to that demonstrated for cCTnC. Synthesis of peptide cIp is as previously described in Chapter III (*33, 36*). For titration, binding and ^{15}N-relaxation measurements of the C-lobe of cTnC in complex with cRP40 (cTnI$_{33\text{-}71}$), a synthetic peptide was synthesized via standard peptide synthetic procedures, with the sequence acetyl-AKKKSKISASRKLQLKTLLLQIAKQELEREAEERRGEK-amide. A phosphorylated (cRP40-Ser41Phos/Ser43Phos) version of this peptide was also synthesized via standard synthetic procedures with an amino acid sequence containing acetyl-AKKKSKI$\underline{S}_p$A$\underline{S}_p$RKLQLKTLLLQIAKQELEREAEERRGEK-amide. Following synthesis, all synthetic peptides were lyophilized twice to remove any residual organic solvents.

V-B. Competition binding studies of cIp and cRP40 peptides for ^{15}N-cCTnC•2Ca^{2+}

13.88mg of ^{15}N-cCTnC was dissolved in 1100*u*L of NMR buffer (pH 6.7), sterile filtered and 525*u*L was separately aliquoted into two clean 5mm NMR tubes. For the first titration series (cRP40 followed by cIp), a total of 12 titration points were acquired when solid lyophilized cRP40 peptide was added directly to one of the NMR tubes containing ^{15}N-cCTnC•2Ca^{2+} in 0.1mg additions for the first 4 additions, followed by 3 additions of 0.25mg, 3 additions of 0.5mg and 2 final additions of 1mg to reach a saturated ^{15}N-cCTnC•2Ca^{2+}•cRP40 complex. Following the final cRP40 titration point, translational (T_1) and transverse (T_2) ^{15}N-NMR relaxation experiments were acquired. 8.38mg of solid, lyophilized cIp peptide was dissolved in 50*u*L of NMR buffer (pH 6.7) and centrifuged in a 1.5mL sterile, 0.22*u*m Nylon spin-filter (Costar, Corning Incorporated) at 3K rpm for 10 minutes to remove all impurities and any undissolved sample. To the sample tube containing ^{15}N-cCTnC•2Ca^{2+}•cRP40, a total of 5 titration points were acquired with additions of 2, 5, 10, 15, and 18*u*L of the cIp solution to create a saturated complex of ^{15}N-cCTnC•2Ca^{2+}•cRP40•cIp. T_1 and T_2 ^{15}N-NMR relaxation experiments were acquired. Both 1D ^{1}H-NMR and 2D ^{1}H/^{15}N-HSQC spectra were acquired at every titration point, with the pH kept constant using small additions of 0.1M HCl when required.

For the second titration series (cIp followed by cRP40), 8.29mg of cIp peptide was dissolved in 50*u*L of NMR buffer (pH 6.7) and was prepared identically to that for the first titration series. The titration was composed of a total of 10 data points in which there was 4 separate additions of 1*u*L of the cIp solution added directly to the other NMR tube containing ^{15}N-cCTnC•2Ca^{2+}, followed by 2 additions of 2*u*L, and single additions of 3, 5, 10 and 24*u*L respectively to create a saturated complex of ^{15}N-cCTnC•2Ca^{2+}•cIp. T_1 and T_2 ^{15}N-NMR relaxation experiments were acquired. A further titration of cRP40 to the ^{15}N-cCTnC•2Ca^{2+}•cIp complex was then preformed with a total of 10 data points in which solid, lyophilized cRP40 peptide was added directly to the NMR tube in 0.1mg additions for the first 3 additions, followed by 3 additions of 0.25mg, 2 additions of 0.5mg and 2 final additions of 1mg. T_1 and T_2 ^{15}N-NMR relaxation experiments were acquired. Both 1D ^{1}H-NMR and 2D ^{1}H/^{15}N-HSQC spectra were acquired at every titration point, with the pH kept constant using small additions of 0.1M HCl when required.

V-C. Titration of ^{15}N-cTnC•3Ca^{2+} with cRP40 and cRP40-S41Phos/S43Phos

Both cRP40 and cRP40-S41Phos/S43Phos peptides are highly soluble in water, which allows the preparation of a 35mM stock solution of cRP40 and a 40mM stock solution of cRP40-S41Phos/S43Phos. A 500*u*L NMR sample of 0.6mM ^{15}N-cTnC•3Ca^{2+} was used for the titration with cRP40 and a 500*u*L NMR sample of 0.7mM ^{15}N-cTnC•3Ca^{2+} was used for the titration with cRP40-S41Phos/S43Phos. Each titration was done by adding consecutive aliquots of 1*u*L of cRP40 or cRP40-S41Phos/S43Phos stock solutions. The sample was mixed thoroughly with each addition. The change in protein concentration due to dilution was taken into account for data analysis. The decrease in pH associated with cRP40 or cRP40-S41Phos/S43Phos addition was adjusted by 1M NaOH to pH 6.7 at every titration point. Both 1D 1H and 2D {^{1}H, ^{15}N}-HSQC NMR spectra were acquired at every titration point.

V-D. NMR Spectroscopy

All NMR spectral data were obtained using a Unity INOVA 500MHz spectrometer at 30°C. All 1D ^{1}H-NMR spectra were acquired using a sweep width of 6000Hz with 32 transients. All 2D {^{1}H, ^{15}N}-HSQC spectra were acquired using the sensitivity-enhanced gradient pulse scheme developed by Lewis E. Kay (*37*, *38*). The ^{1}H and ^{15}N sweep widths were 6000Hz or 7000Hz and 1500Hz, with 16 transients and 128 increments (1024 points in the directly detected dimension), respectively. Spectral processing of all {^{1}H, ^{15}N}-HSQC was performed using the software package NMRPipe by Frank Delaglio (*39*), and spectral viewing, assignments and relaxation fit using the software package NMRView by Bruce Johnson (*40*). All samples were referenced directly (indirectly for ^{15}N dimension) to a 0.2mM internal standard of 2,2-Dimethyl-2-silapentane-5-sulfonate (Cambridge Isotope Laboratories).

V-E. ^{15}N-NMR T_1 and T_2 relaxation parameters

All relaxation data were acquired using an INOVA 500MHz NMR spectrometer at 30°C. The delay times used for ^{15}N-T_1 relaxation measurements were set to 11.1, 55.5, 122.1, 199.8, 277.5, 388.5, 499.5, 666, 888, and 1100 ms respectively. The delay times used for ^{15}N-T_2 relaxation measurements were set to 16.61 × *n* ms (where *n* = 1, ..., 11). All other parameters were set equal to the {^{1}H, ^{15}N}-HSQC parameters previously listed.

RESULTS

V-A Binding studies of cIp and cRP40 on C-lobe of cTnC

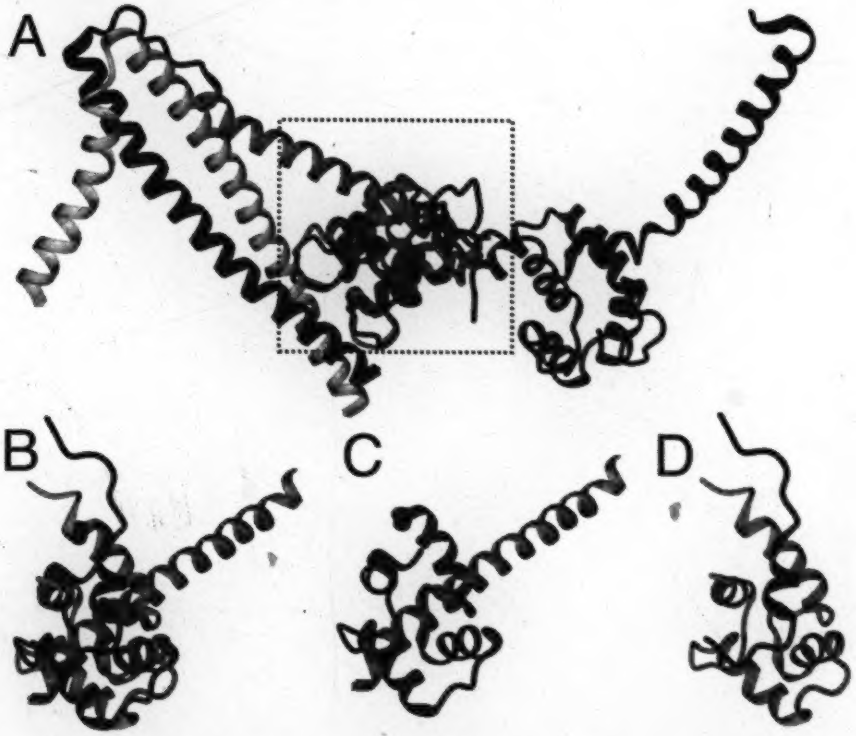


Figure V-1. **Steric clash of the inhibitory and cRP40 regions of cTnI.** (A) C-lobe superimposition of the NMR binary complex (green) and the core crystal structure (red). (B) Re-orientation of the C-lobe to display (C) the crystal structure cRP40 region and (D) the NMR structure of the inhibitory region.

The cRP40 and clp domains share a common binding site on the C-lobe, such that superimposition of the two solved structures reveals a steric clash (Figure V-1). To ascertain the binding parameters of these two domains in the presence of one another, competition-binding titrations of the C-lobe with both clp and cRP40 were monitored by {^{1}H, ^{15}N}-HSQC NMR spectroscopy. The titration in Figure V-2 shows a primary titration of clp peptide to ^{15}N-cCTnC (Figure V-2A), followed by a secondary titration of the ^{15}N-cCTnC•2Ca^{2+}•clp complex with cRP40 (Figure V-2B). Large chemical shift perturbations for both additions indicate that a separate binding event is taking place during the addition of each respective peptide to the C-lobe. The binding of clp to the C-lobe indicates that it is within the fast exchange NMR timescale, however the binding of cRP40 is within the slow exchange NMR timescale.

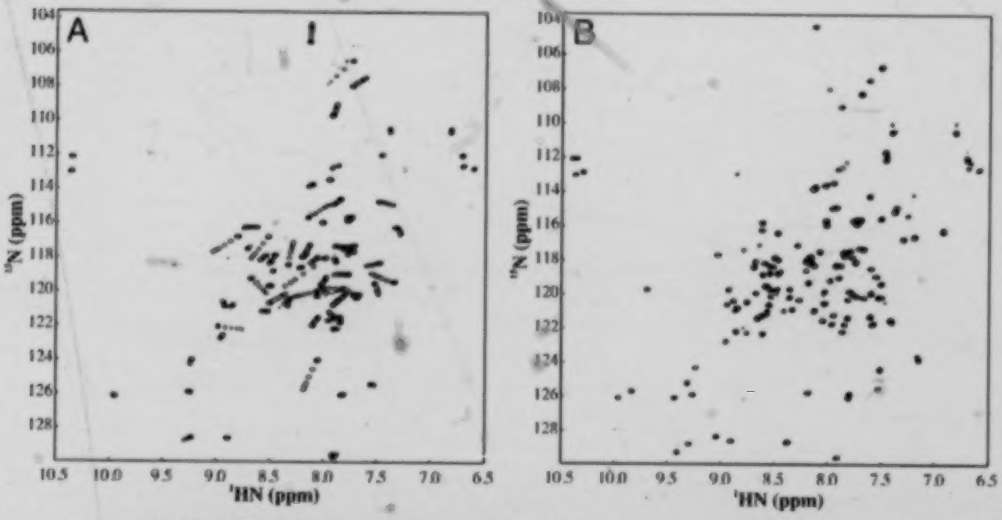


Figure V-2. **^{15}N-HSQC plots of ^{15}N-cCTnC•Ca^{2+} titration with clp and cRP40.** ^{15}N-cCTnC•Ca^{2+} (black dots) titrated first with clp peptide (A – clear red circles), followed by cRP40 peptide (B – blue dots).

A complementary titration to that seen in Figure V-2 was required to definitively determine the binding properties of the two peptides in the presence of one another. The primary titration observed in Figure V-3 shows cRP40 binding to ^{15}N-cCTnC (Figure V-3A) followed by the secondary addition of clp to the ^{15}N-cCTnC•2Ca^{2+}•cRP40 complex (Figure V-3B). The secondary addition of clp to the ^{15}N-cCTnC•2Ca^{2+}•cRP40 complex to make a ^{15}N-cCTnC•2Ca^{2+}•cRP40•cIp complex resulted in minute perturbations of the C-lobe backbone resonances. As observed for the titration in Figure V-2, the binding of cRP40 was within the slow exchange limit and the binding of cIp was within the fast exchange limit of the NMR timescale limit for both titrations.

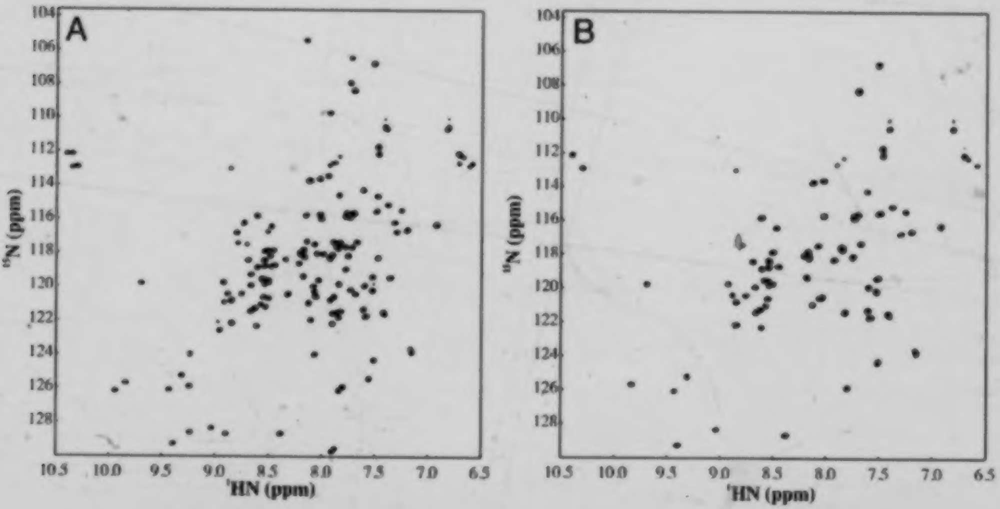


Figure V-3. **^{15}N-HSQC plots of ^{15}N-cCTnC•Ca^{2+} titration with cRP40 and cIp.** ^{15}N-cCTnC•Ca^{2+} (black dots) titrated first with cRP40 peptide (A – blue dots), followed by cIp peptide (B – red dots).

The {^{1}H, ^{15}N}-HSQC spectra of the final bound species of the C-lobe in the presence of both cIp and cRP40 were undistinguishable whether the titration began with cIp or cRP40 first (Figure V-2B - blue dots, and V-3B - red dots), yielding the conclusion that the final complex was not dependent on the titration order of the peptides (i.e. cRP40 does not require cIp to bind correctly and vis-versa). The final titration plots of the saturated complexes resulted in chemical shifts that were nearly identical to that for only the cRP40 region bound to the C-lobe (Figure V-3A - blue dots). This result was identical to that previously reported for a similar skeletal isoform titration (*19*). Similar conclusions that were proposed for the skeletal system can be drawn for the cardiac system, in that it appears as though cRP40 is capable of completely displacing the cIp peptide completely off of the C-lobe. This indicates that although the two peptides share a common binding region, cRP40 dominates when the two peptides are in competition for binding.

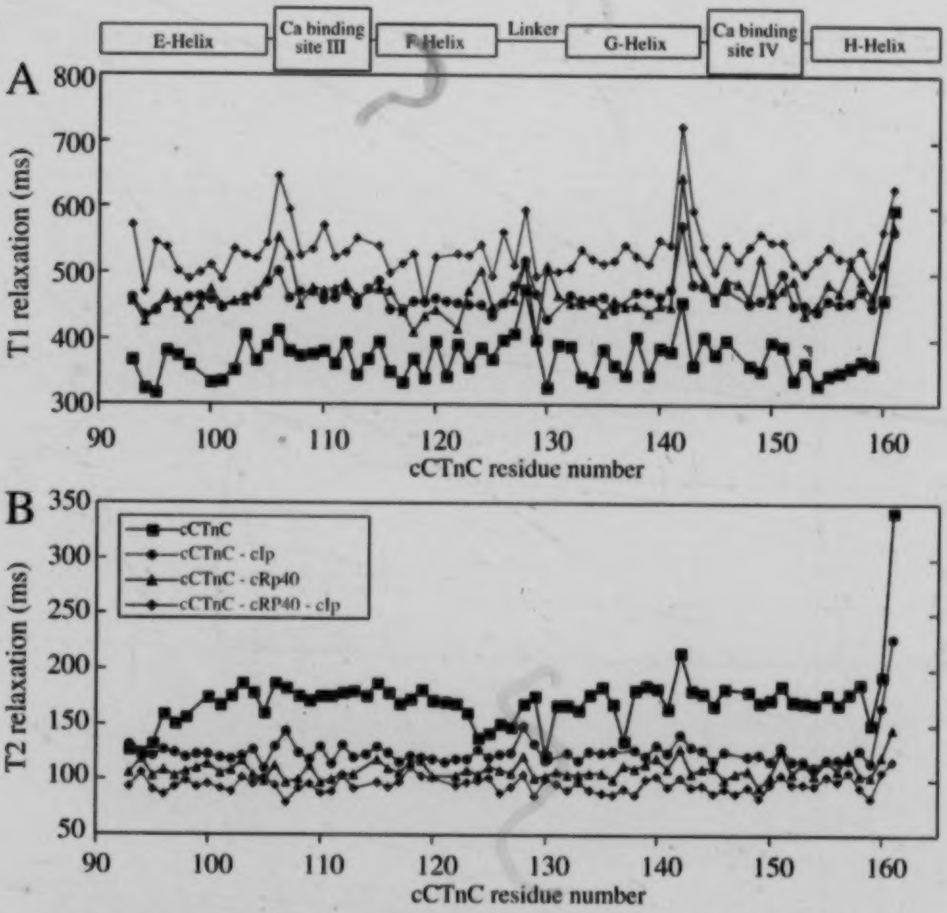


Figure V-4. **^{15}N-NMR backbone relaxation of the C-lobe of cTnC in complex with cTnI peptides**. The secondary structural elements of the C-lobe backbone are diagramed above the T_1 values. The color legend present in the T_2 graph depicts the 4 investigated troponin protein complexes for both A and B.

To verify the ability of cRP40 to completely displace cIp off the C-lobe, I acquired 4 sets of ^{15}N-NMR relaxation data at 500MHz of ^{15}N-cCTnC•2Ca^{2+} in the unbound form and bound to cIp alone, cRP40 alone, and both cIp and cRP40 together (Figure V-4). The predicted increases in molecular weight of the complexes due to peptide binding were mirrored by the transient increases in T_1 values and decreases in T_2

values for the C-lobe resonances. The complex of ^{15}N-cCTnC•2Ca^{2+}•cRP40•cIp revealed higher average T_1 and lower average T_2 values than that for the ^{15}N-cCTnC•2Ca^{2+}•cRP40 complex, indicating that cIp and cRP40 are both bound to the C-lobe. Further additions of either cIp and/or cRP40 in an attempt to super-saturate the ^{15}N-cCTnC•2Ca^{2+}•cRP40•cIp complex with either peptide did not change relaxation values, indicating that the complex is present in a 1:1:1 ratios of C-lobe•cRP40•cIp, and that there is no non-specific binding of either of the peptides to the C-lobe.

Figure V-5 demonstrates the linear relationship of ^{15}N-R2 (1/T_2) versus molecular weight for the 4 data sets obtained in this study, compared against other relaxation values obtained from other protein complexes within my laboratory (i.e. sCTnC•2Ca^{2+}, sCTnC•2Ca^{2+}•sRP40, sNTnC, cNTnC, cNTnC•Ca^{2+}) all obtained under similar conditions (ie. 30°C, spectrometer frequency of 500MHz). Each individual protein complex has an associated average R2 relaxation value that was calculated using only those individual backbone resonances, which were within one standard deviation of the mean relaxation value. All are globular proteins and assumed to behave as isotropically tumbling bodies in solution.

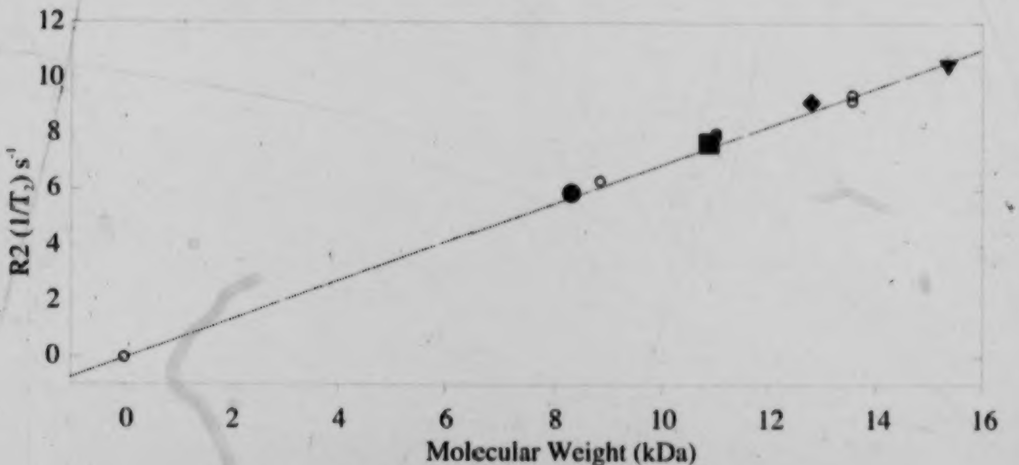


Figure V-5. **Linear relationship of ^{15}N-NMR R2 relaxation values versus molecular weight for various troponin protein complexes (● - ^{15}N-cCTnC•2Ca^{2+}, ■ - ^{15}N-cCTnC•2Ca^{2+}•cIp, ◆ - ^{15}N-cCTnC•2Ca^{2+}•cRP40, ▼ - ^{15}N-cCTnC•2Ca^{2+}•cIp•cRP40, O - other troponin protein complexes).**

V-B. Effect of Ser41/Ser43 phosphorylation on the interaction of cRP40 with cTnC•3Ca^{2+}

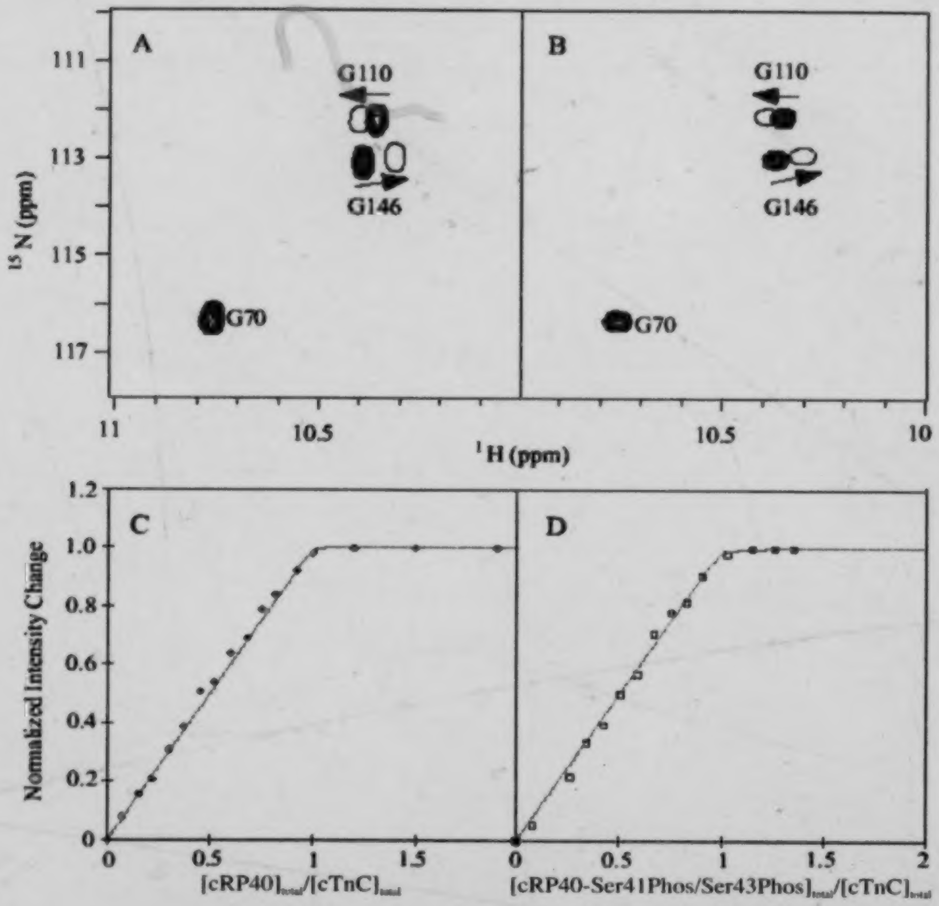


Figure V-6. **Titration of ^{15}N-cTnC•3Ca^{2+} with cRP40 and cRP40-Ser41Phos/Ser43Phos.** In the expanded contour plots of the 2D {^{1}H, ^{15}N}-HSQC spectra shown in (A) and (B), the cross-peaks corresponding to free cTnC•3Ca^{2+} are shown as multiple contours, whereas the peaks corresponding to the protein•peptide complexes are shown as single contours. The titration curves for cRP40 and cRP40-Ser41Phos/Ser43Phos binding are shown in (C) and (D), respectively. Each curve is normalized according to $(I_{obs} - I_{initial})/(I_{final} - I_{initial})$, where I represents the cross peak intensity changes. A solid line represents the best fit to the data.

Previously, it has been shown that cRP40 binds cCTnC•2Ca^{2+} in a 1:1 stoichiometry to form a tight cCTnC•2Ca^{2+}•cRP40 complex ($K_D \leq 1 u$M) and the interaction of cRP40 with cCTnC•2Ca^{2+} occurs with slow exchange kinetics on the NMR time scale (*7*). A similar binding behavior of cRP40 for cTnC•3Ca^{2+} indicates that the interaction of this region of cTnI binds specifically to the C-lobe of cTnC•3Ca^{2+} regardless of the presence of the N-lobe (*14*). Here I compare the binding behavior of cRP40 and cRP40-Ser41Phos/Ser43Phos. Figure V-6A shows a superimposition of an expanded region of the 2D {^{1}H, ^{15}N}-HSQC NMR spectra of cTnC•3Ca^{2+} and cTnC•3Ca^{2+}•cRP40 complex and Figure V-6B shows a superimposition of an expanded region of the 2D {^{1}H, ^{15}N}-HSQC NMR spectra of cTnC•3Ca^{2+} and cTnC•3Ca^{2+}•cRP40-Ser41Phos/Ser43Phos complex. The 2D {^{1}H, ^{15}N}-HSQC NMR spectrum of cTnC•3Ca^{2+} has been previously assigned and was used as a guide in assigning some of the residues in the cRP40 or cRP40-Ser41Phos/Ser43Phos bound states (*36*). In both cases, the C-lobe residues underwent significant perturbations, with several C-lobe residues (i.e. Gly110/Gly146) displayed in Figures V-6A and V-6B as examples. As the titration progressed, the resonance peaks corresponding to cTnC•3Ca^{2+} became less intense while those corresponding to the cTnC•3Ca^{2+}•cRP40 or cTnC•3Ca^{2+}•cRP40-Ser41Phos/Ser43Phos complex grew. When the protein:peptide ratio reached 1:1, all cross peaks corresponding to cTnC•3Ca^{2+} completely disappeared, while those corresponding to the complex attained maximum intensity. When the intensity changes as a function of peptide to protein ratios were plotted, the curve levels off after protein:peptide ratio reaches 1:1. Figures V-6C and V-6D reveal the respective binding curves of cRP40 and cRP40-Ser41Phos/Ser43Phos for cTnC•3Ca^{2+}. In both cases, curve fitting yielded a dissociation constant $K_D \leq 1uM$. It is important to point out that since the K_D values of $\leq 1uM$ are far less than the concentrations of protein used in the NMR experiments, an affinity less than 1*u*M becomes difficult to determine from a titration curve, (*i.e.* it is hard to distinguish between K_Ds of 1*u*M, 0.1*u*M, and 0.01*u*M). Thus, my results indicate that the phosphorylation of Ser41/Ser43 does not destabilize the tight association of cRP40 with cTnC•3Ca^{2+} as the K_D is still equal or below 1*u*M (indicative of high affinity between ligand and protein). Close inspection of the cRP40-cTnC interaction as per the crystal structure cTnC•3Ca^{2+}•cTnI$_{31\text{-}210}$•cTnT$_{183\text{-}288}$ (*3*), shows that the cRP40 region forms a long α-helix, interacting primarily with the C-lobe of cTnC•3Ca^{2+} via multiple polar and van der Waals interactions, similar to the earlier structure of sTnC•2Ca^{2+}•sTnI$_{1\text{-}47}$ (*24*).

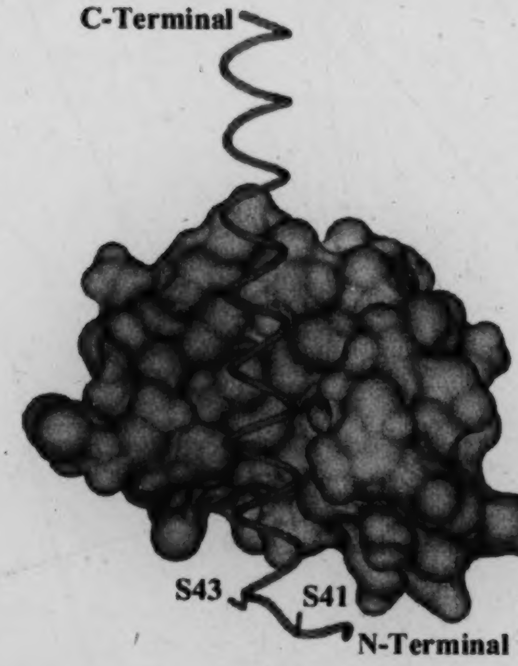


Figure V-7. **Ribbon and surface representations of the structure of cCTnC•2Ca^{2+}•cRP40.** Coordinates derived from the cTnC•3Ca^{2+}•cTnI$_{31-210}$•cTnT$_{183-288}$ complex (PDB accession code 1J1E) (*3*). The positions of Ser41 and Ser43 phosphorylation are indicated.

As shown in Figure V-7, the α-helix enters the C-lobe of cTnC hydrophobic cleft near helices G/H and exits near helices E/F and the Ser41/Ser43 phosphorylation sites are located approximately in the beginning of the α-helix with almost no direct contacts with the protein hydrophobic cleft. Thus, it is unlikely that the introduction of two negative charged phosphate groups at Ser41 and/or Ser43 would exert much influence on the extensive contacts between the long α-helix of cRP40 and the C-lobe of cTnC. Interestingly, in both the sTnC•2Ca^{2+}•sTnI$_{1-47}$ and cTnC•3Ca^{2+}•cTnI$_{31-210}$•cTnT$_{183-288}$ structures, TnC adopts a collapsed conformation as such the N-terminus of the cRP40 α-helix makes several weak contacts with the N-lobe of cTnC. This is reflected in the observation that some of the N-lobe peaks (i.e. Gly42, Val72, Val79) underwent slight chemical shift changes in the titration of cRP40-Ser41Phos/Ser43Phos to cTnC•3Ca^{2+} (data not shown). It is possible that the two negative charged phosphate groups at Ser41Phos/Ser43Phos caused minor conformational alterations in the N-lobe of cTnC.

DISCUSSION

In this chapter, I have expanded my efforts to probe the binding effects of the inhibitory region of cTnI with that of the C-lobe of cTnC. The steric clash observed between the inhibitory region and the cRP40 regions when superimposition of the binary NMR complex with the core crystal structure indicates that the two domains occupy a common binding face on the C-lobe (Figure V-1). However the ^{15}N-NMR relaxation data reported for the ^{15}N-cIp peptide bound to the C-lobe in Chapter III, as well as the electrostatic analysis reported in Chapter IV, firmly establishes that the inhibitory region binds to the C-lobe in a manner shown in the binary NMR structure. The core crystal structure refutes the NMR binary structure as the cRP40 region was shown to be bound within the binding site for the inhibitory region. The study reported in this chapter was therefore proposed to address this conflict and yield a solution to the problem.

Previous {^{1}H, ^{15}N}-HSQC NMR results reported for the skeletal isoform of the inhibitory region (sIp), RP40 region (sRP40) and C-lobe (sCTnC) indicated that sRP40 would completely displace sIp off the sCTnC domain (*19*). To investigate if the reported skeletal conclusions were valid for the cardiac isoform, and indeed if cRP40 would displace cIp off the C-lobe of cTnC, I repeated the skeletal binding titration experiments with the cardiac isoforms. In addition to the titrations, I included ^{15}N-NMR T_1 and T_2 relaxation data acquired at specific points of saturation to determine binding effects of the two peptides on the C-lobe. For the titration of the C-lobe first with cIp (Figure V-2A), the resulting chemical shifts were identical to those previously reported in Chapter III for the wild-type peptide (*33*). The secondary titration of cRP40 to the cCTnC•cIp complex (Figure V-2B) revealed final C-lobe chemical shifts that we nearly identical to that for just cCTnC-cRP40 alone (Figure V-3A), suggesting that cRP40 is capable of completely displacing cIp off the C-lobe. These results mirror and extend those of Mercier, who reached the same conclusion for the skeletal isoform (*19*). The addition of cIp to a cCTnC-cRP40 complex (Figure V-3B) induced very minimal C-lobe chemical shift perturbations, indicating that once the cRP40 domain is bound to the C-lobe, cIp is no longer able to bind to the C-lobe.

^{15}N-NMR relaxation data analysis reveals a significantly different conclusion than that reached from {^{1}H, ^{15}N}-HSQC titration data (Figure V-4). The T_1 and T_2 values for the cCTnC•cRP40•cIp complex are significantly altered than that for the cCTnC•cRP40 and/or the cCTnC•cIp complexes alone. This relaxation data reveals that the conclusions based solely on the chemical shift mapping data by {^{1}H, ^{15}N}-HSQC NMR techniques were pre-mature and incorrect. The inhibitory and cRP40 regions are both capable of simultaneous binding to the C-lobe. The simple plot of R2 versus total molecular weight (Figure V-6) reveals a linear relationship for the various C-lobe•peptide troponin complexes, when measured against various other protein complexes studied within our laboratory. The molecular weight of the cCTnC•cRP40 complex is increasing by the additional mass of the cIp peptide (2.6 kDa) upon binding, indicating that the binding event is capable of a 1:1:1 complex of C-lobe, cRP40 and cIp respectively.

The observation that the addition of cIp to the cCTnC•cRP40 complex induced minimal chemical shift perturbations to the C-lobe (Figure V-3B) is very significant in relation to the relaxation data and the previous electrostatic binding results for cIp. The sensitivity of the {^{1}H, ^{15}N}-HSQC backbone amide bond chemical shift to the chemical environment around it reveals that for cIp binding to the C-lobe in the presence of cRP40, the binding event is not near the hydrophobic cleft of the C-lobe, but instead far away from the backbone. Instead it is probable that salt-bridging electrostatic interactions of the side-chains are anchoring the cIp region onto the cCTnC•cRP40 complex with non-perturbing results to the C-lobe amide chemical shift values.

This data suggests that troponin domain orientation(s) might be similar to the 'Lindhout' model discussed in Chapter III and may be a more realistic representation of the physiological complex. The Lindhout model is re-displayed in Figure V-8 for viewer reference. The ability for the C-lobe to cooperatively bind both cRP40 and cIp domains leads to the conclusion that the crystal structure may contain only one possible 'low temperature' conformation of the complex. Therefore the intrinsic flexibility of the troponin complex might be much higher than previous estimates may have predicted. Perhaps the true structure of the troponin complex is more of a mosaic, where there may be some movement possible about the D/E linker of cTnC as it shuffles between the calcium bound and unbound forms during cardiac rhythm. In addition, the ability of cTnI to toggle between cTnC and its actin binding site(s) during the cardiac cycle, dictates that molecular motion is not only an *intrinsic*, but is more a *necessary* component of cardiac contraction.

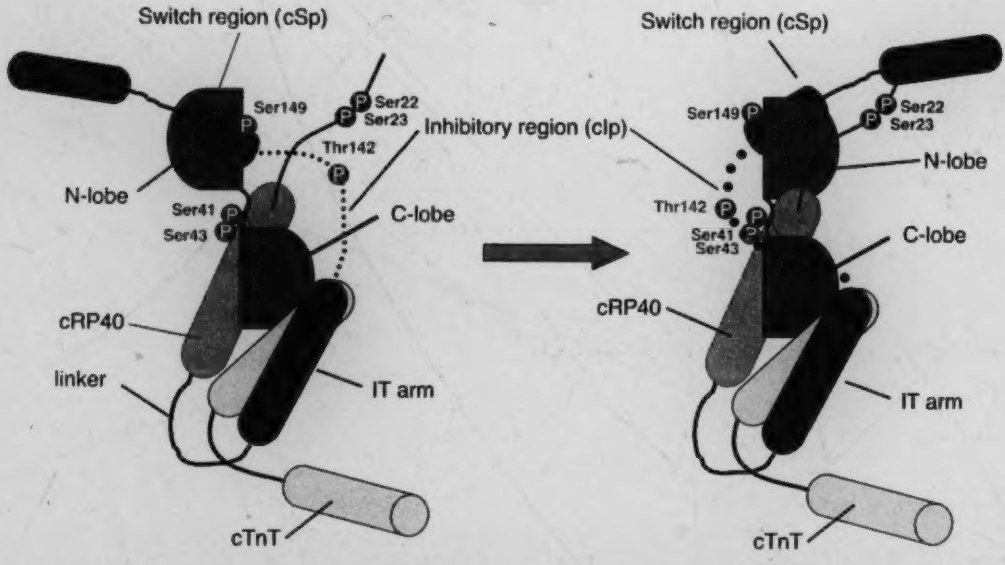

Switch region (cSp)
Ser149
Ser22
Ser23
Inhibitory region (cIp)
Thr142
N-lobe
Ser41
Ser43
C-lobe
cRP40
linker
IT arm
cTnT
Switch region (cSp)
Ser149
Ser22
Ser23
N-lobe
Thr142
Ser41
Ser43
C-lobe
cRP40
IT arm
cTnT

Figure V-8. **'Lindhout' model of the 'core' cardiac troponin complex**. Cartoon model of the Takeda structure (left), where a re-orientation of the N-lobe about the D/E linker (right) would allow the inhibitory region to come into contact with the C-lobe. Sites of cTnI phosphorylation by PKC (Red - Ser41/Ser43/Thr142), PAK (Grey - Ser149) and PKA (Purple – Ser22/Ser23) are indicated.

Interestingly, the helical region of the inhibitory region in my binary NMR structure is the only part of cIp that is in direct steric conflict with the Takeda crystal structure (1J1E.pdb); not the extended, basic C-terminal tail composing residues Arg140-Arg147 (see Figure V-1). It has been proposed that the electrostatic attraction between the C-lobe and the cIp region is the dominant force directing association (see Chapter IV). Therefore the steric clash observed for residues Phe132-Phe138 of the inhibitory region becomes somewhat irrelevant as these residues are implicated to be part of a coiled-coil region with cTnT, and when complexed, are not found in the area of hydrophobic groove of the C-lobe. Unfortunately the Takeda cardiac crystal structure was unable to visualize these interactions with an inhibitory region. However, the recent crystal structure of the skeletal isoform of the troponin complex by Vinogradova (1YTZ.pdb) was successful in elucidating these important contacts (*1*). In this structure, the C-terminal tail of the inhibitory region assumes an extended conformation, while making numerous electrostatic interactions with residues along the D/E helix (Figure V-9). This structure is remarkably consistent with the cardiac model proposed in Figure V-8 (and the model proposed in Figure III-16, Chapter III, page 98); the only difference being a non-flexible helix between the D/E helical regions.

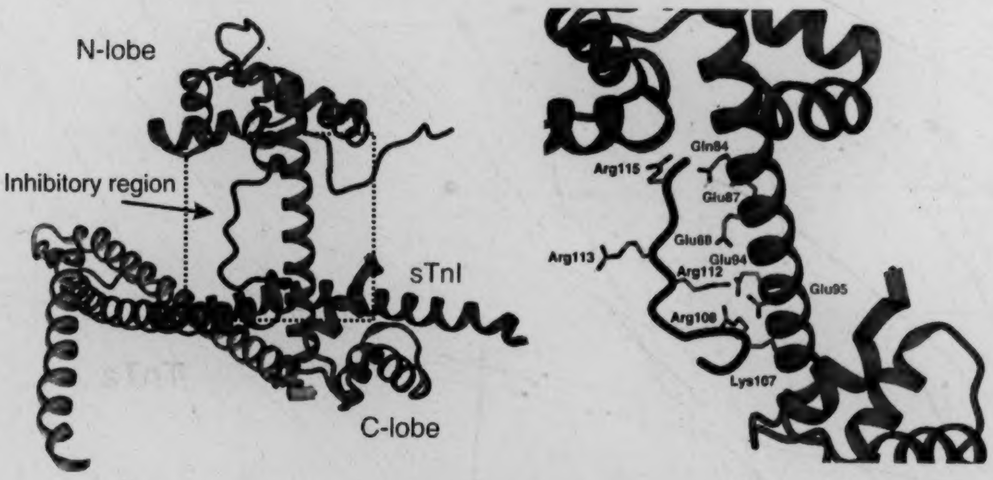


Figure V-9. **Crystal structure of the skeletal troponin complex**. The crystal structure is displayed to the left as a ribbon diagram, with sTnC in green, sTnT in yellow and sTnI in cyan and the inhibitory region in purple (*1*). An expanded view of residues involved in electrostatic interactions is diagrammed to the right. Skeletal number of residues is indicated (1YTZ.pdb).

Although the scope of this crystal structure is outside the discussion for this chapter, an expanded discussion regarding the NMR binary structure, the cardiac crystal structure and the skeletal crystal structure will be investigated in Chapters VII & VIII of this thesis, with a focused discussion regarding the positioning of the inhibitory region and the physiological/functional implications of these interactions for both the cardiac and other isoforms.

Three sites of cTnI phosphorylation have been studied by NMR spectroscopy in Chapters IV & V that includes three specific PKC sites at residues Ser41, Ser43 and Thr142. Although Ser41/Ser43 are sequentially distant from Thr142, the three PKC sites are in close proximity with one another in the troponin complex, and the proposed model (Figure V-8). Thus, it is not surprising that all three sites are phosphorylated upon the activation of PKC in cardiocytes. The relative contribution of the phosphorylation of three sites to the accelerated relaxation and decreased acto-myosin ATPase activity has been the focus of several recent studies. For example, Burkart reported that phosphorylation of Ser41/Ser43 dominates regulation of the level of maximum tension, while phosphorylation of Thr142 appears to be required for regulation of thin filament sliding speed (*41*). The data presented in this thesis has demonstrated that it is the phosphorylation of Thr142, not Ser41/Ser43, which plays a major role in destabilizing the cTnI•cTnC interaction, that subsequently leads to diminished acto-myosin ATPase activity and depressed contractility. Since Ser41/Ser43 are located in the structural domain of the troponin complex based on the proposed model (*3*), the phosphorylation of Ser41/Ser43 may be involved in the global conformational changes of troponin in the thin filament rather than in the cTnI•cTnC interface. In addition to the known regulatory PKC sites within cTnI, there are other sites of phosphorylation such as Ser149 by p21-activated kinase (PAK) and Ser22/Ser23 by PKA. As Ser149 of cTnI is located in the center of the hydrophobic core of cNTnC•Ca^{2+} (*6*), introduction of a negatively charged phosphate group in this region of cTnI would certainly destabilize its binding to cNTnC•Ca^{2+} and result in reduced affinities (see Chapter VI for an expanded study). Although Ser22/Ser23 appear at present to be located in an area void of structural details, it is difficult to comment on their specific roles at the molecular level. However insights from the core troponin structure indicates that these might possess a spatial placement close to the Ca^{2+}-sensitizing N-lobe, which may aid in explanations of potential influences that these residues might elicit on contractile activity (see Figure V-8). As well, there may be potential interactions of Ser22/Ser23 with thin filament proteins such as actin, tropomyosin and/or the 2nd cTnI actin-binding site (cAp - $cTnI_{163-190}$) in directing physiological regulatory effects. Given that the skeletal isoform does not contain this 33 residue-extension N-terminal PKA phosphorylation domain, this may play a key role in shaping the differing structural, regulatory and physiological roles of cardiac tissues. Nonetheless, the dramatic affinity reduction of cTnI for cTnC by the phosphorylation of Thr142 strongly suggests that kinases such as PKC (and/or PAK) may modulate cardiac muscle contraction by modifying myofilament protein-protein interaction, contributing to depressed contractility in certain cases and compromised relaxation in other.

REFERENCES

1. Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) *Proc Natl Acad Sci U S A.*
2. Sia, S. K., Li, M. X., Spyracopoulos, L., Gagné, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) *J. Biol. Chem. 272*, 18216-18221.
3. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
4. Takeda, S., and Maeda, Y. (2003) *Seikagaku 75*, 1540-5.
5. Li, Y., Love, M. L., Putkey, J. A., and Cohen, C. (2000) *Proc. Natl. Acad. Sci. U. S. A. 97*, 5140-5.
6. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.
7. Wang, X., Li, M. X., Spyracopoulos, L., Beier, N., Chandra, M., Solaro, R. J., and Sykes, B. D. (2001) *J Biol Chem 276*, 25456-66.
8. Wang, X., Li, M. X., and Sykes, B. D. (2002) *J Biol Chem 277*, 31124-33.
9. Lindhout, D. A., and Sykes, B. D. (2003) *J Biol Chem 278*, 27024-34.
10. Dvoretsky, A., Abusamhadneh, E. M., Howarth, J. W., and Rosevear, P. R. (2002) *J Biol Chem 277*, 38565-70.
11. Tripet, B., De Crescenzo, G., Grothe, S., O'Connor-McCourt, M., and Hodges, R. (2002) *J Mol Biol 323*, 345.
12. Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) *J Mol Biol 271*, 728-50.
13. Blumenschein, T. M., Tripet, B. P., Hodges, R. S., and Sykes, B. D. (2001) *J Biol Chem 276*, 36606-12.
14. Li, M. X., Saude, E. J., Wang, X., Pearlstone, J. R., Smillie, L. B., and Sykes, B. D. (2002) *Eur Biophys J 31*, 245-56.
15. Li, M. X., Spyracopoulos, L., Beier, N., Putkey, J. A., and Sykes, B. D. (2000) *Biochemistry 39*, 8782-8790.
16. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1998) *Biophys. J. 74*, A51.
17. Li, Z., Gergely, J., and Tao, T. (2001) *Biophys J 81*, 321-33.
18. Mercier, P., Ferguson, R. E., Irving, M., Corrie, J. E., Trentham, D. R., and Sykes, B. D. (2003) *Biochemistry 42*, 4333-48.
19. Mercier, P., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 2902-11.
20. Mercier, P., Spyracopoulos, L., and Sykes, B. D. (2001) *Biochemistry 40*, 10063-77.
21. Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) *FEBS Lett 453*, 107-12.
22. Finley, N., Dvoretsky, A., and Rosevear, P. R. (2000) *J Mol Cell Cardiol 32*, 1439-46.
23. Finley, N. L., Howarth, J. W., and Rosevear, P. R. (2004) *Biochemistry 43*, 11371-9.
24. Vassylyev, D. G., Takeda, S., Wakatsuki, S., Maeda, K., and Maeda, Y. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 4847-4852.
25. Brown, L. J., Sale, K. L., Hills, R., Rouviere, C., Song, L., Zhang, X., and Fajer, P. G. (2002) *Proc Natl Acad Sci U S A 99*, 12765-70.

26. Campbell, A. P., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) *Biochim* ~~*Biophys Acta 1160*~~, 35-54.
27. Campbell, A. P., and Sykes, B. D. (1991) *J. Mol. Biol. 222*, 405-421.
28. Hernandez, G., Blumenthal, D. K., Kennedy, M. A., Unkefer, C. J., and Trewhella, J. (1999) *Biochemistry 38*, 6911-7.
29. Tung, C. S., Wall, M. E., Gallagher, S. C., and Trewhella, J. (2000) *Protein Sci 9*, 1312-26.
30. Heller, W. T., Abusamhadneh, E., Finley, N., Rosevear, P. R., and Trewhella, J. (2002) *Biochemistry 41*, 15654-63.
31. Talbot, J. A., and Hodges, R. S. (1979) *J. Biol. Chem. 254*, 3720-3723.
32. Sheldahl, C., Xing, J., Dong, W. J., Harvey, S. C., and Cheung, H. C. (2003) *Biophys J 84*, 1057-64.
33. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.
34. Chandra, M., Dong, W. J., Pan, B. S., Cheung, H. C., and Solaro, R. J. (1997) *Biochemistry 36*, 13305-13311.
35. Li, M. X., Gagné, S. M., Tsuda, S., Kay, C. M., Smillie, L. B., and Sykes, B. D. (1995) *Biochemistry 34*, 8330-8340.
36. Li, M. X., Wang, X., Lindhout, D. A., Buscemi, N., Van Eyk, J. E., and Sykes, B. D. (2003) *Biochemistry 42*, 14460-8.
37. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.
38. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.
39. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
40. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603-614.
41. Burkart, E. M., Sumandea, M. P., Kobayashi, T., Nili, M., Martin, A. F., Homsher, E., and Solaro, R. J. (2003) *J Biol Chem 278*, 11265-72.

CHAPTER VI

Troponin interactions with the N-lobe of cardiac troponin C

OVERVIEW

This chapter continues my investigation into the interactions of troponin I, however it will expand the focus from the C-lobe (Chapters III, IV & V) to include the N-lobe of cardiac troponin C. In the first half of this chapter, interactions of the inhibitory region will be explored. The second half of this chapter will address familial hypertrophic cardiomyopathy and mutational (R161W) and p21-activated kinase phosphorylation (Ser149) events occurring within the switch region of troponin I. The FHC and phosphorylation studies are intended to complement the previous results observed in Chapter IV, which were reported for the inhibitory region with the C-lobe (i.e. R144G and Thr142). The studies involving the inhibitory region was performed solely by myself, and work involving the switch region was performed primarily by Xu Wang and myself. Additionally, Monica X. Li, Nina Buscemi, Jennifer E. Van Eyk, and Brian D. Sykes were contributors to the FHC and phosphorylation data reported within this chapter. The majority of the data presented in this chapter has been published in two journals, namely the journal *Biochemistry* on December 16th, 2003 [*Biochemistry*. 2003 Dec 16;**42(49)**:14460-8.] and *Proceedings for the International School of Structural Biology and Magnetic Resonance, 6th Course on Protein Dynamics, Function and Design* on February 1st, 2005 [*Structure, Dynamics and Function of Biological Macromolecules and Assemblies.* **V364.** ISO Press, Amsterdam, The Netherlands, Editor J. D. Puglisi.].

Calcium induced interactions of troponin C with troponin I on the thin filament of heart tissue plays a central role in muscle contraction. The inhibitory region of troponin I is known to interact with the C-lobe of troponin C and the switch region of troponin I is know to interact with the N-lobe of troponin C respectively. Using ^{15}N-HSQC NMR spectroscopy, I have been successful in monitoring the competition for the N-terminal lobe of cardiac troponin C by both the inhibitory and switch regions of troponin I. The inhibitory region is displaced off the N-lobe by the switch region, yet the inhibitory region can initiate a partially "opening" of the N-lobe. In addition, I will explore the effects of mutation (R161W) and phosphorylation (Ser149) on the binding of the switch region to the N-lobe of cTnC. The binding affinities measured for the N-lobe with cSp will be examined against the reported results of the C-lobe with cIp (Chapter IV of this thesis), and conclusions based upon similarities/differences of the N-lobe versus C-lobe during cTnI binding will be addressed. Combined with the phosphorylation data observed for the cIp and cRP40 domains, I have addressed the concept of phosphorylation as a regulatory and functionally significant event, in the fine-tuning of cardiac contraction.

INTRODUCTION

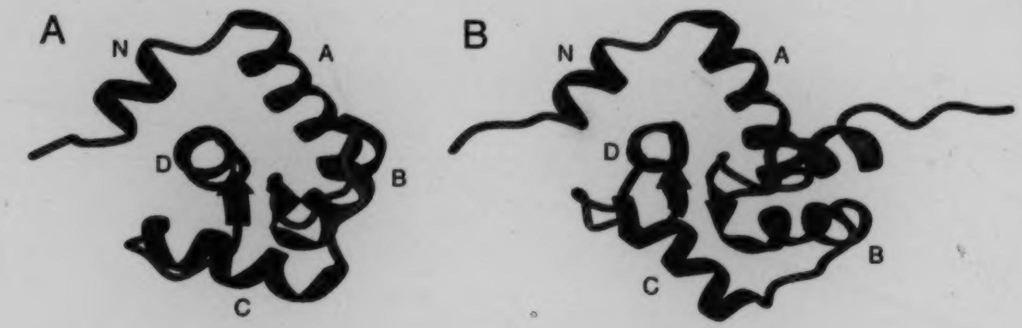


Figure VI-1. **Ribbon representation of cNTnC in (A) the calcium bound state and (B) the cSp bound state.** cNTnC helices are colored in black and cSp is colored in grey. Reorientation of helices B and C reveals an "opening" of the N-lobe upon Sp binding. Figures A & B were created using the PDB coordinates (A) 1SPY and (B) 1MXL.

The previously stated work by Head and Perry in their early biochemical and proteolytic analysis of TnI, defined a 17 residue section ($cTnI_{147-163}$), which is C-terminal to the inhibitory region (cIp), that contained unique binding properties separate from that of the inhibitory and other TnI regions (*1*, *2*). This 17 amino acid region, which will be referred to as the 'switch' region (cSp), interacts in a calcium-dependant manner with the N-lobe of cardiac troponin C (cTnC). Structural studies on the N-terminal regulatory lobe of the cardiac isoform of TnC has demonstrated that binding of a single calcium ion in site II is sufficient to induce a conformational change from a "closed" to an "open" state upon the binding of cSp (*3*). The calcium saturated N-lobe possesses an increased affinity for cSp, which is able to bind within a hydrophobic cleft on the surface of the molecule (*4, 5*). This 'opening' of the N-lobe is characterized by significant reorientations of the B and C helices relative to the N, A and D helices of cNTnC (*6, 7*). cSp adopts helical a helical conformation upon binding, making hydrophobic interactions with the N-lobe and fits within the A and B helices (Figure VI-1).

As it has now been firmly established in previous chapters: contraction and relaxation of cardiac tissue is mediated by the *whole* troponin complex. Therefore limiting my studies solely to a specific region of cTnI (i.e. the inhibitory region), would not allow me to obtain an overall picture of what is truly happening on the thin filament during the calcium-signaling event. I felt that I must expand my studies to include functionally important areas of cTnI such as the switch region (and additional segments such as cRP40 – see Chapter VI), to establish a relationship of the troponin complex with the thin filament, and with cTnC. As cTnC initiates heart contraction upon the binding of calcium, structural changes are initiated in the N- and/or C-lobes, cTnI senses the structural changes within cTnC and in the presence of tropomyosin moves off the actin-myosin binding site to make numerous interactions with TnC, allowing myosin-ATPase activity with actin to generate the power stroke (*8*).

Numerous structural, as well as biophysical studies, have been performed to better understand the molecular mechanisms for muscle contraction (*9-15*). It has been established that the switch region does not display affinity for actin/tropomyosin during the contraction cycle: it is *unique* when *compared* to the *structural and functional* duties of the inhibitory region. The binding of the switch region to the N-lobe is sufficient to initiate cardiac contraction, and in effect, is able to 'pull' the inhibitory region off its binding partner actin along with it as it binds to cTnC. This binding event alone is sufficient to garner it attention for study in cardiac contraction. However, the specific interactions and affinities of the switch region and inhibitory region individually for the N-lobe have never been established or reported. Does the inhibitory region possess an affinity for the N-lobe? Should an affinity be established, what would be the effects of the addition of the switch region to this complex? Does the inhibitory region play a role in the calcium-induced movement of cTnI to cTnC during contraction? In attempting to elucidate an understanding to these questions, I have studied this system using ^{15}N heteronuclear single quantum coherence (HSQC) NMR spectroscopy. I have investigated the potential for competition for cNTnC by both cIp and cSp regions of cTnI in the calcium saturated state. Unexpectedly, the cIp region binds to the N-lobe; yet is displaced by the stronger binding of the cSp peptide. The affinity of cIp to cNTnC suggests that it may have sufficient binding to initiate a partial "opening" of the N-lobe in the absence of cSp.

I have previously demonstrated in Chapter IV that cTnI and cTnI-cTnC interactions are subject to modifications by intracellular signaling molecules, which play vital roles in mediating the myofilament activation. Phosphorylation of cTnI appears to be of particular importance in modulating myofilament activity. Two serine residues (Ser22 and Ser23), that are substrates for protein kinase A (PKA) lie within the N-terminal extension of cTnI (for a review, see (*16*)). cTnI is also sensitive to protein kinase C (PKC) phosphorylation at Ser41, Ser43, and Thr142. The effect of cTnI phosphorylation by those kinases is primarily a downward modulation of cardiac contractility, either by decreased actomyosin ATPase activity, which is mediated by PKC, or by enhanced Ca^{2+} dissociation from cTnC and reduced myofilament Ca^{2+} sensitivity, mediated by PKA. Recently, a study has shown that p21-activated kinase (PAK) increases the Ca^{2+}-sensitivity of triton-skinned cardiac muscle fiber bundles via a mechanism potentially involving the phosphorylation of cTnI at Ser149 (*17*), representing a novel cTnI phosphorylation site in addition to the PKA and PKC sites. cTnI mutation is also implicated in various cardiac diseases. For example, familial hypertrophic cardiomyopathy (FHC), one of the most frequently occurring inherited cardiac disorders has been identified to be a sarcomere disease (for reviews, see (*18, 19*))) and is believed to be caused by mutations in certain contractile protein genes, including cTnI. Among those, R144G is located in the key inhibitory region (see Chapter IV) and R161W is located in the important switch region of cTnI.

In order to understand how alterations in cTnI are involved in disease, it is important to elucidate the effects of cTnI modifications on the interaction with cTnC. Although many functional and animal model experiments have revealed the differences between native and modified cTnI that may underline the pathogenesis of cardiac disease (*20-26*), there is little data available concerning the structural consequences of cTnI modification with regard to the interaction of cTnI with cTnC or cTnI with actin-tropomyosin, especially on the combined effects of FHC mutation with phosphorylation of cTnI. In this study, I have utilized 2D $\{^1H, {}^{15}N\}$-HSQC NMR spectroscopy to elucidate the binding the cSp segment to the N-lobe of cTnC with R161W in the mutated and/or phosphorylated states. The affinity of cSp for cNTnC•Ca^{2+} was reduced by the modifications. An important and novel finding of this work was that while the FHC mutation R144G enhances the effect of Thr142 phosphorylation on the interaction of cIp and cCTnC•$2Ca^{2+}$ (see Chapter IV of this thesis), the FHC mutation R161W suppresses the effect of Ser149 phosphorylation on the interaction of cSp and cNTnC•Ca^{2+}, demonstrating linkages between FHC mutation and phosphorylation of cTnI. The observed alterations corroborate well with the structural data. The results suggest that those in the key functional cTnI region (cIp and cSp) have a pronounced effect on the interaction of cTnI and cTnC, which may correlate with the altered myofilament function and cardiac muscle contraction under pathophysiological conditions.

EXPERIMENTAL PROCEDURES

VI-A. Preparation of cNTnC protein and synthetic peptides

The engineering of the cNTnC-A_{cys} vector (1-89) was as described by Wang et al. (*27*) and expression and purification of ^{15}N-cNTnC is identical as for that of the C-lobe (*28-30*), which is as described previously in chapter III of this thesis. The synthetic peptides cSp ($cTnI_{147-163}$); acetyl-**RISADAMMQALLGARAK**-amide, cSp-S149Phos ($cTnI_{147-163}$); acetyl-**RIS$_p$ADAMMQALLGARAK**-amide, cSp-R161W ($cTnI_{147-163}$); acetyl-

RISADAMMQALLGAWAK-amide, cSp-R161W-S149Phos (cTnI$_{147-163}$); acetyl-RIS$_p$ADAMMQALLGAWAK-amide and peptide cIp (cTnI$_{128-147}$); acetyl-TQKIFDLRGKFKRPTLRRVR-amide, were synthesized as previously described in chapter III of this thesis (*31*).

VI-B. Titration of cIp with the N-lobe of cTnC

2.12mg of ^{15}N-cNTnC was dissolved in 600*u*L of NMR buffer (100mM KCl, 10mM imidazole, 15mM Ca^{2+}, 0.1mM DSS, pH=6.7, 90%/10% H_2O/D_2O). 500*u*L of the resulting solution was added to an NMR tube. 2.63mg of cIp was dissolved in 15*u*L of NMR buffer to which 1*u*L was added to the NMR tube for the first 4 titration points, followed by an addition of 2*u*L for the next 2 titration points, followed by an addition of 3*u*L and a final addition of 4*u*L for ensure complete cNTnC saturation of cIp. {^{1}H, ^{15}N}-HSQC spectra were acquired at each titration point to monitor chemical shift changes of cNTnC backbone resonances upon cIp binding.

VI-C. Titration of cSp with cNTnC•cIp complex

To determine if cSp could displace cIp off the N-lobe, A total of 9 titration points were performed with cSp additions of solid lyophilized powder directly to the NMR tube in the following amounts: 0.14mg, 0.22mg, 0.34mg, 0.22mg, 0.24mg, 0.35mg, 0.50mg, 1.13mg and 1.03mg. cSp peptide is slightly soluble in aqueous solution and tends to form a gel at high concentrations (>50mM), likely due to aggregation. Thus, no stock peptide solution was prepared; instead, solid peptide was added directly to the NMR tube at every titration point. {^{1}H, ^{15}N}-HSQC spectra were acquired at each titration point to monitor chemical shift changes of cNTnC backbone resonances upon cSp binding. The change in sample pH during the titration was negligible and corrected for with small volumes of 0.1M NaOH

VI-D. Titration of cSp with the N-lobe of cTnC

13.3mg of ^{15}N-cNTnC was dissolved in 600*u*L of NMR buffer and 500*u*L was added to an NMR tube. Solid, lyophilized cSp powder was added directly to the NMR tube in exactly 1.0mg amounts for 6 additions, followed by a final addition of 2.0mg and {^{1}H, ^{15}N}-HSQC spectra were acquired at each titration point to monitor chemical shift changes of cNTnC backbone resonances upon cSp binding.

VI-E. Titration of cIp with cNTnC•cSp complex

9.28mg of lyophilized cIp powder was dissolved in 50*u*L of NMR buffer to which 2*u*L was added to the NMR tube for the first 2 titration points, followed by additions of 3*u*L for the next 2 titration points. When no visible changes in the backbone chemical shift resonances for cNTnC were observed upon the first 4 additions of cIp to the cNTnC•cSp solution, 2 final additions of 20*u*L of cIp were made in an attempt to saturate the cNTnC•cSp complex. ^{15}N-HSQC spectra were acquired at each titration point to monitor chemical shift changes of cNTnC backbone resonances upon cIp additions. The change in sample pH during the titration was negligible and corrected for with small volumes of 0.1M NaOH.

VI-F. Titration of ^{15}N-cNTnC with cSp, cSp-S149Phos, cSp-R161W, and cSp-R161W-S149Phos

A 500*u*L (0.6mM) sample of ^{15}N-cNTnC•Ca^{2+} was used for the cSp titration, a 500*u*L (0.5mM) sample was used for the cSp-S149Phos titration, a 500*u*L (0.7mM) sample was used for the cSp-R161W titration, and a 500*u*L (0.6mM) sample was used for the cSp-R161W-S149Phos titration. The sample was mixed thoroughly with each addition. 2*u*L of sample in duplicate were taken at every titration point for amino acid analysis to determine the peptide/protein ratios. Changes in pH associated with cSp additions were compensated by adjusting to pH 6.7 at every titration point. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC NMR spectra were acquired at every titration point.

VI-G. NMR spectroscopy

For titrations of the N-lobe involving cIp and cSp peptides, NMR spectra were recorded on a Varian INOVA 500MHz NMR spectrometer at 30°C. 1D ^{1}H-NMR spectra were acquired with a sweep width of 6000Hz or 7000Hz with 128 transients. All 2D {^{1}H, ^{15}N}-HSQC spectra were acquired using the sensitivity-enhanced gradient pulse scheme developed by Lewis E. Kay and co-workers (*32, 33*). The ^{1}H and ^{15}N sweep widths were 6000Hz or 7000Hz and 1500Hz, with 16 transients and 128 increments (1024 points in the directly detected dimension), respectively. All 2D NMR spectra were zero-filled, back predicted using linear prediction and processed using nmrPipe (*34*) and analyzed on NmrView (*35*) for assignment purposes.

RESULTS

VI-A. Effects of cIp and cSp binding on the N-lobe of cTnC

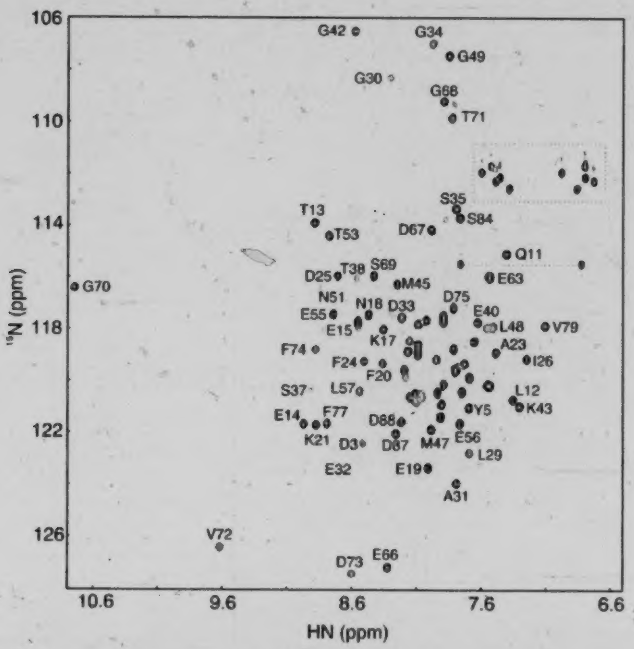


Figure VI-2. **^{15}N-HSQC spectra of cNTnC-A$_{cys}$.** Residue assignments of various cross-peaks in the calcium saturated state are as indicated. Amide (CONH$_2$) side-chain moieties of Asn and Gln residues are indicated by a dotted line.

The {^{1}H, ^{15}N}-HSQC NMR spectra of calcium saturated cNTnC-Acys revealed a soluble, monomeric domain with sufficient spectral dispersion to allow for complete backbone assignments of all residues (Figure VI-2). It has been well established by numerous independent structural studies that cNTnC binds cSp upon a calcium influx, with cSp adopting a helical conformation upon binding (see Figure VI-1), sitting on the hydrophobic cleft upon an "opening" of the B & C helices relative to the N, A & D helices of cNTnC (*4, 5, 7, 27, 36*).

The titration of cNTnC with cIp peptide indicates that cIp binding to the N-lobe occurs in a linear fashion, with increasing amounts of cIp causing increased chemical shift changes of the cNTnC backbone. The addition of cSp to the cNTnC•cIp complex caused significant chemical shift changes in a greater number of the N-lobe backbone resonances, indicating a broader binding region for the cSp region on cNTnC (Figure VI-3). Analysis of the data suggests one of two possibilities regarding the two part titration: (1) cIp and cSp both bind cooperatively to the N-lobe, or (2) cSp displaces cIp off the N-lobe as the binding affinity of cSp for cNTnC is higher than that for cIp.

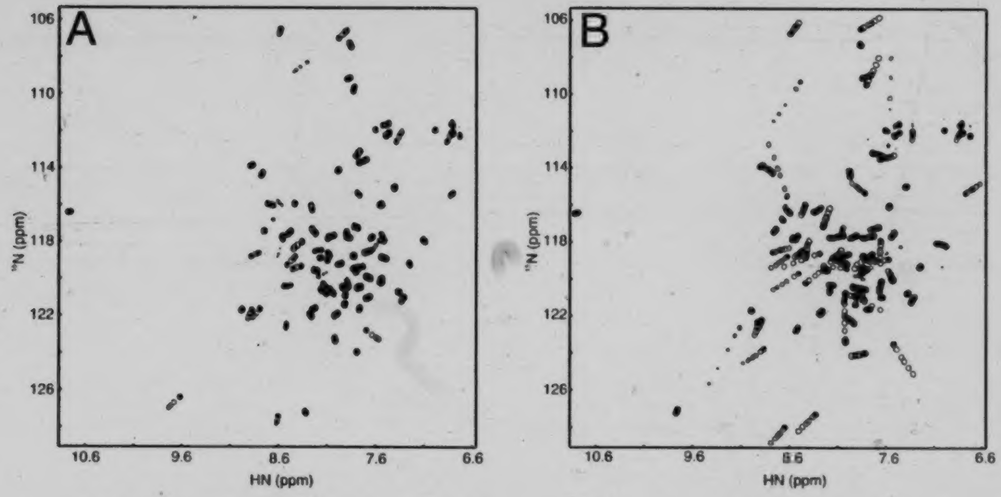


Figure VI-3. **cIp-cSp titration of ^{15}N-cNTnC-A$_{cys}$.** The addition of (A) cIp peptide, followed by the addition of (B) cSp peptide, are shown in the {^{1}H, ^{15}N}-HSQC NMR spectra above. The starting spectra are indicated by a dark cross-peak (●) and subsequent titration points by light cross-peaks (○).

To determine which of the two hypotheses was correct, reversing the addition order of the peptides in the titration (cSp first, cIp second) was performed to obtain an answer. The titration with cSp first produced pronounced backbone chemical shift perturbations within cNTnC, which upon saturation had final chemical shift resonances identical to that of cNTnC when previously titrated with cIp and then cSp (Figure VI-3B). Analysis of this data reveals that cSp peptide competes for the binding surface on the N-lobe with cIp and will displace cIp from the N-lobe. Not surprising, the addition of cIp to the cNTnC•cSp complex yielded negligible chemical shift changes within the backbone resonances of cNTnC (Figure VI-4).

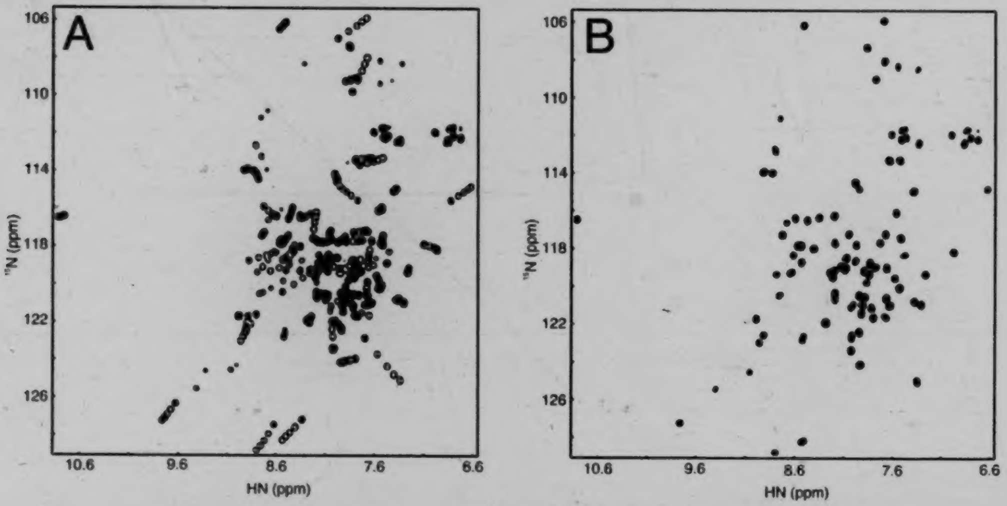


Figure VI-4. **cSp-cIp titration of ^{15}N-cNTnC-A$_{cys}$.** The addition of (A) cSp peptide, followed by the addition of (B) cIp peptide, are shown in the {^{1}H, ^{15}N}-HSQC NMR spectra above. The starting spectra are indicated by a dark cross-peak (●) and subsequent titration points by light cross-peaks (○).

To understand the interactions of cIp and cSp with the N-lobe, I investigated the published structure by Li and Spyracopoulos of cSp bound to cNTnC to gain insights into specific protein-peptide interactions of the bound complex (*7*). Close inspection of the structure of cNTnC•cSp complex reveals that Arg147 of the cSp peptide is within close proximity to the N-lobe in the bound complex (Figure VI-5). Within this complex, Arg147 adopts a position over the B and C helices of the N-lobe. Analysis of the amino acid sequences of cSp and cIp reveals that Arg147 is the one common residue shared between both cSp and cIp synthetic peptides, with Arg147 being the C-terminal residue on cIp and the N-terminal residue on cSp. It is thus proposed that the binding of cIp to cNTnC occurs at the C-terminal end of cIp, in that Arg147 and possibly Arg145 and Val146 of cIp may play a role in binding.

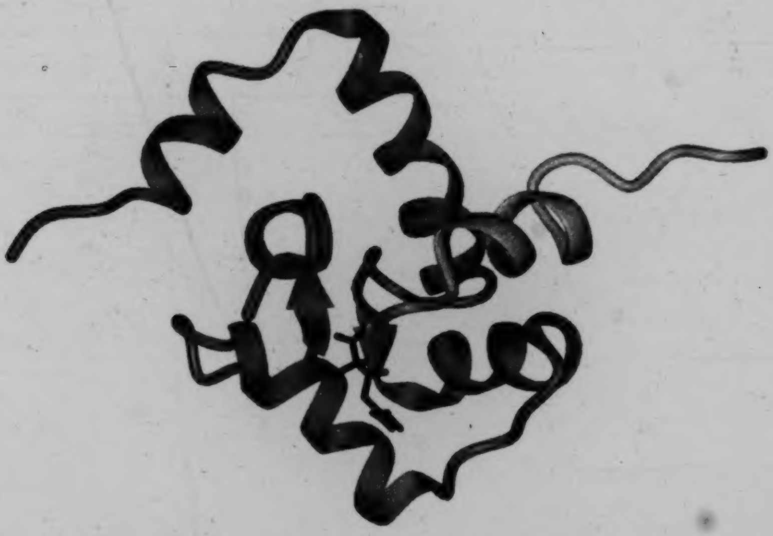


Figure VI-5. **Ribbon diagram of calcium saturated cNTnC bound to cSp.** cNTnC helices are colored in green and cSp is colored in yellow. Arg147 is colored red at the N-terminal end of cSp and lies over the B and C helices of the N-lobe. This diagram was created using the PDB coordinates 1MXL.

To better understand the local effects of the binding of cIp on the N-lobe, I looked at the chemical shift mapping of cNTnC backbone resonances upon cIp addition (Figure VI-3A). The {^{1}H, ^{15}N}-HSQC spectral titration revealed numerous shifted cross-peaks, which indicated the localized region of putative binding. These shifted resonances were mapped onto the Connolly surface of the published structure of cNTnC in the calcium-saturated state (3). The changes in chemical shifts of the cNTnC backbone amide resonances were measured as per equation [1]:

$$\Delta\delta(\mathrm{ppm}) = [\,(\Delta\delta H)^2 + (\Delta\delta N)^2\,]^{1/2} \qquad [1]$$

Residues that underwent amide chemical shift changes greater than one standard deviation of the average of the whole were colored in black to illustrate the localized effects of cIp binding on the N-lobe. Amide resonances that underwent lesser chemical shift changes were colored according to a color gradient of white to black, with lighter regions undergoing little or no chemical shift changes upon cIp binding. The localized effects of cIp binding to the N-lobe by chemical shift mapping indicate that the region near the A/B helices and the hinge between them underwent the largest localized changes in chemical shift perturbation (Figure VI-6).

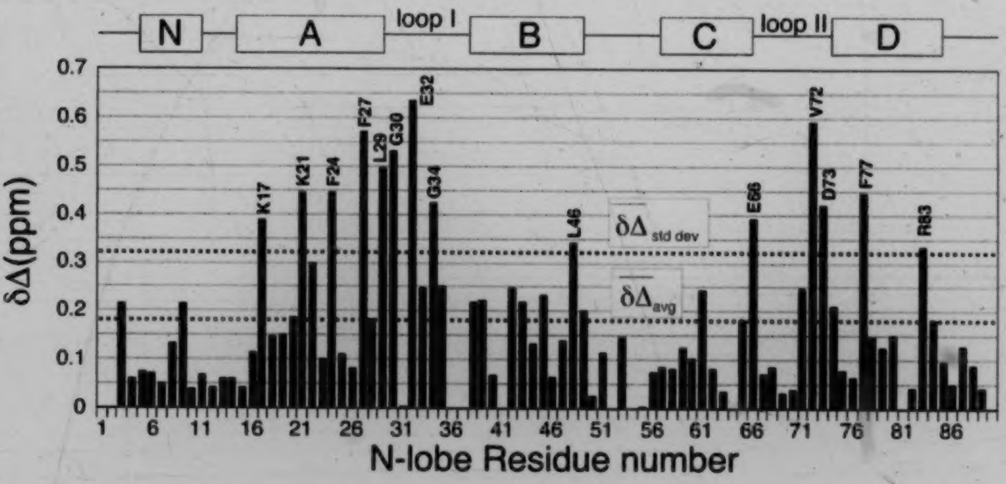


Figure VI-6. **N-lobe induced chemical shifts upon cIp binding**. Induced backbone chemical shift changes for the N-lobe in the presence of cIp. Residues undergoing greater than one standard deviation of the average are indicated.

This effect is similar to that of the chemical shift perturbation results of the binding of cSp to the N-lobe, as the structural opening of the N-lobe in response to the cSp peptide causes changes in chemical shifts near or at the A/B-hinge region. This observed effect describes the power of chemical shift mapping methodology as it is able to deduce structural changes in an area of a protein, which is not involved directly in the binding event. In this case, the largest induced chemical shift changes are on the back-side of the binding face of the protein, which strongly suggests that the cIp peptide evokes an opening of the N-lobe, although the magnitude of the shifts are much lessened that that for cSp binding. The resulting changes in chemical shift were recorded for each residue in the titration and plotted onto the surface diagram of cNTnC (Figure VI-7). Therefore a partial opening of the N-lobe upon binding cIp is plausible, as cIp does not possess all the necessary hydrophobic residues occurring within cSp to illicit a full opening.

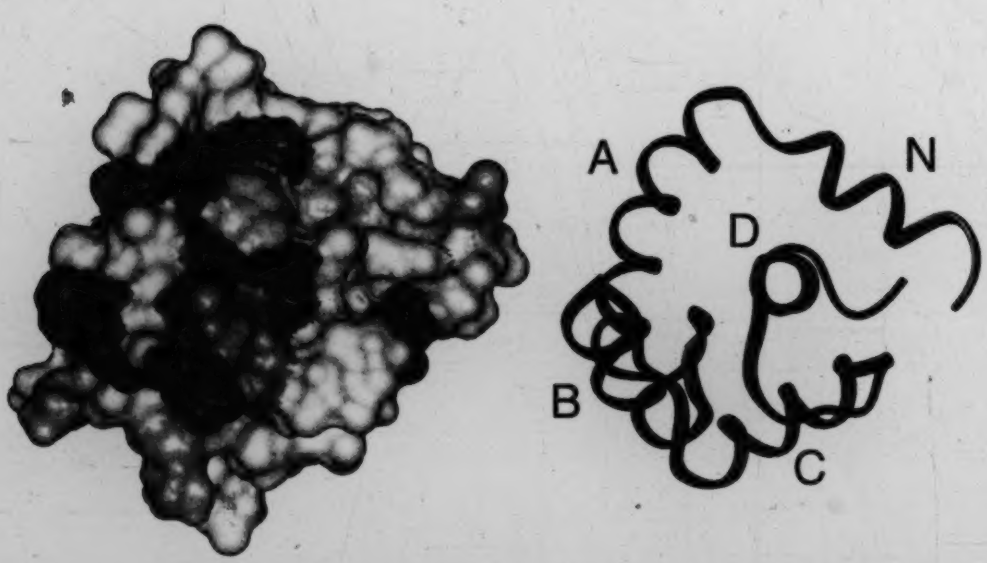


Figure VI-7. **Chemical shift map of the effects of cIp binding on the surface of calcium saturated cNTnC.** Ribbon diagram at right defines the orientation of the surface plot. Darker red indicates areas of higher backbone chemical shift perturbations on the cNTnC backbone upon cIp binding. The surface diagram was generated using INSIGHT II (version 98 Accelrys Inc.), using the PDB coordinates 1SPY.

VI-B. Effect of Ser149 phosphorylation and R161W mutation on the interaction of cSp with cNTnC•Ca^{2+}

The 2D {^{1}H, ^{15}N}-HSQC NMR spectrum of cNTnC•Ca^{2+} was used as a starting point to monitor the induced chemical shift changes of the cSp peptides (see Figure VI-2). The induced spectral changes of ^{15}N-cNTnC•Ca^{2+} in response to binding are very similar to that reported previously for the wild-type cSp peptide (7). The titrations of ^{15}N-cNTnC•Ca^{2+} with cSp-S149Phos, cSp-R161W, and cSp-R161W-S149Phos, respectively, are shown in Figure VI-7. Comparison of the 2D {^{1}H, ^{15}N}-HSQC NMR spectral changes induced by the three cSp peptides, when compared to the wild-type cSp peptide, reveals that all four peptides induce similar pattern of backbone amide resonance changes of cNTnC•Ca^{2+}. Typical examples are G30, G34, and G42 (Figure VI-8A). In all cases, chemical shift changes fall into the fast exchange limit on the NMR time scale and the linear movement of the resonance peaks indicates that the binding occurs with a 1:1 stoichiometry.

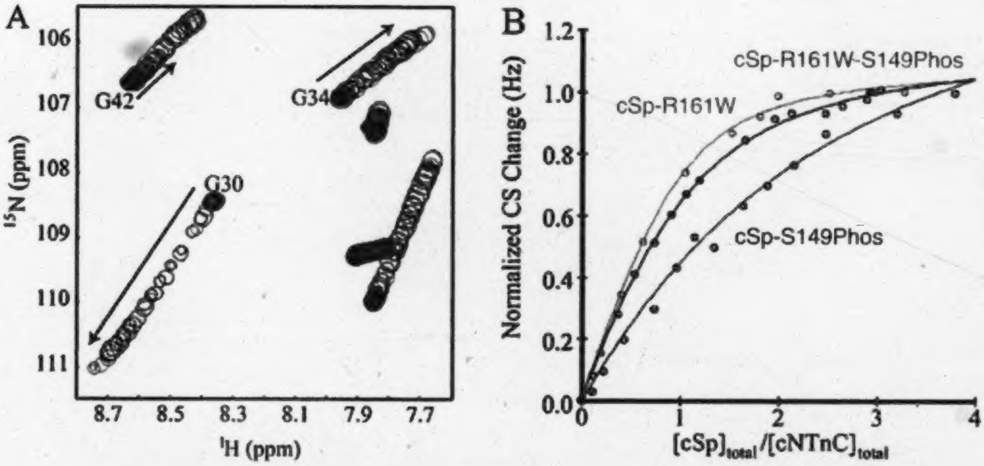


Figure VI-8. **Titration of ^{15}N-cNTnC•Ca^{2+} with cSp-S149Phos, cSp-R161W, and cSp-R161W-S149Phos.** (A) Expanded regions of the 2D {^{1}H, ^{15}N}-HSQC NMR spectra of cNTnC•Ca^{2+} (●) at various peptide additions are superimposed (○), showing the progressive shifts of peaks with increasing peptide concentrations. (B) Titration curves representing the average of all residues followed in panel A. The curves are normalized according to $(\delta_{obs} - \delta_{initial})/(\delta_{complex} - \delta_{initial})$. A solid line represents the best fit to the data.

Resonances that underwent backbone amide ^{1}HN and/or ^{15}N chemical shift changes during titrations were followed to monitor peptide binding to protein. For each titration, the peptide induced chemical shift changes of individual amide resonance as a function of the peptide/protein ratio was plotted, and the average curves (except for cSp, which is near-identical to that published previously; (*7*)) are shown in Figure VI-8B. The normalized chemical shift data were fit to Equation [2] to obtain a dissociation constant (K_D) for each individual peptide (denoted as *) titration. The experimentally determined dissociation constants (K_D) were 130 ± 20 *u*M for cSp, 1.25 ± 0.2 mM for cSp-S149Phos, 176 ± 33 *u*M for cSp-R161W, and 394 ± 32 *u*M for cSp-R161W-S149Phos. This corresponds to an affinity reduction of ~10-fold by the Ser149 phosphorylation, ~1.4-fold by the mutation R161W, and ~4-fold by a combination of Ser149 phosphorylation and mutation R161W.

$$cCTnC \bullet 2Ca^{2+} + cSp^{*} \leftrightarrow cCTnC \bullet 2Ca^{2+} \bullet cSp^{*} \quad [2]$$

The normalized induced chemical shift changes of the N-lobe backbone in response to the binding of the four individual cSp peptides were calculated in a fashion identical to that shown in Equation [1], and plotted against the protein sequence (Figure VI-9A). The important point of this plot is that the patterns induced by the binding of each of the four peptides are very similar, although to a slightly different extent. Although most of the residues throughout the sequence are more or less affected by peptide binding, 26 residues undergo significant ($\Delta\delta ppm \geq 1.5$) peptide-induced chemical shift perturbations. The mapping of those residues on the structure of cNTnC in the $cNTnC \bullet Ca^{2+} \bullet cSp$ complex (Figure VI-9B) reveals that the peptide-induced chemical shift perturbations arise primarily from the A, B, and C-helices and the β-sheet regions. This is not surprising considering that these regions are directly involved in the binding of wild-type cSp and thus undergo large conformational changes in the cSp-induced "closed" to "open" structural transition of the N-lobe.

The fact that these residues are identical in all four cases indicates that neither the Ser149 phosphorylation, the mutation R161W, nor the combination of both, abolish the ability of cSp to induce a "closed" to "open" structural transition in $cNTnC \bullet Ca^{2+}$. It is important to point out that when the 26 residues were mapped to the surface of cNTnC in the $cNTnC \bullet Ca^{2+} \bullet cSp$ complex, it did not reveal a binding patch matching the hydrophobic surface as shown in Figure VI-9C. This is probably because the major conformational perturbation caused by cSp is on the "closed" to "open" structural transition, but not on the binding surface of cNTnC.

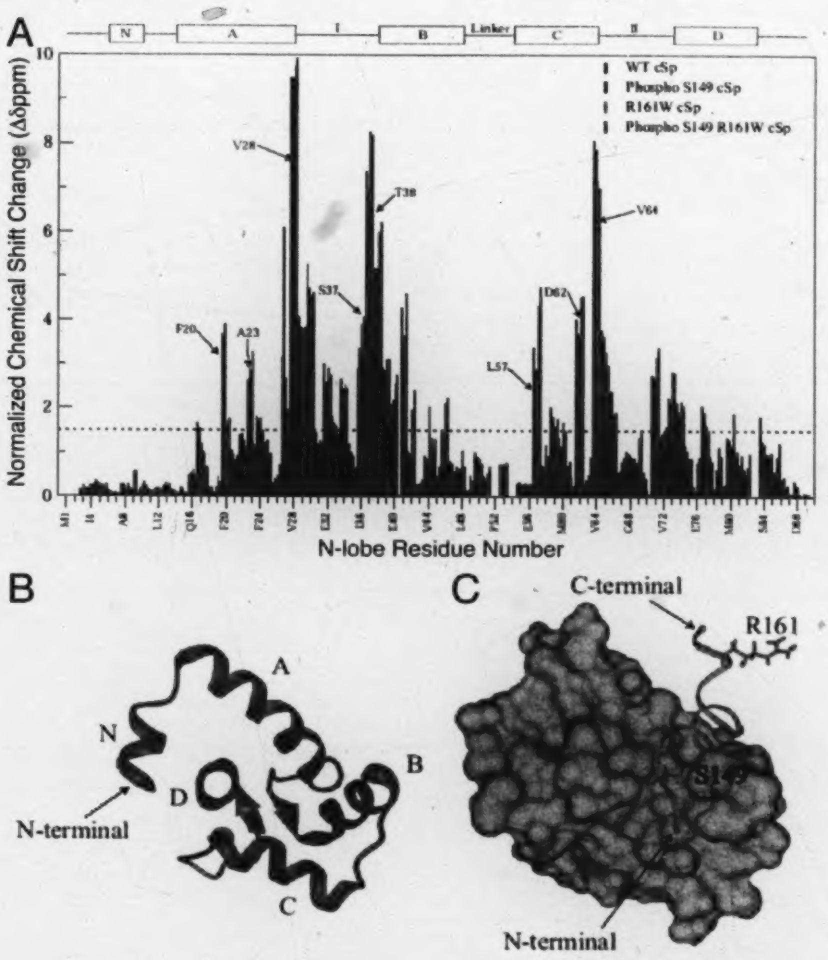


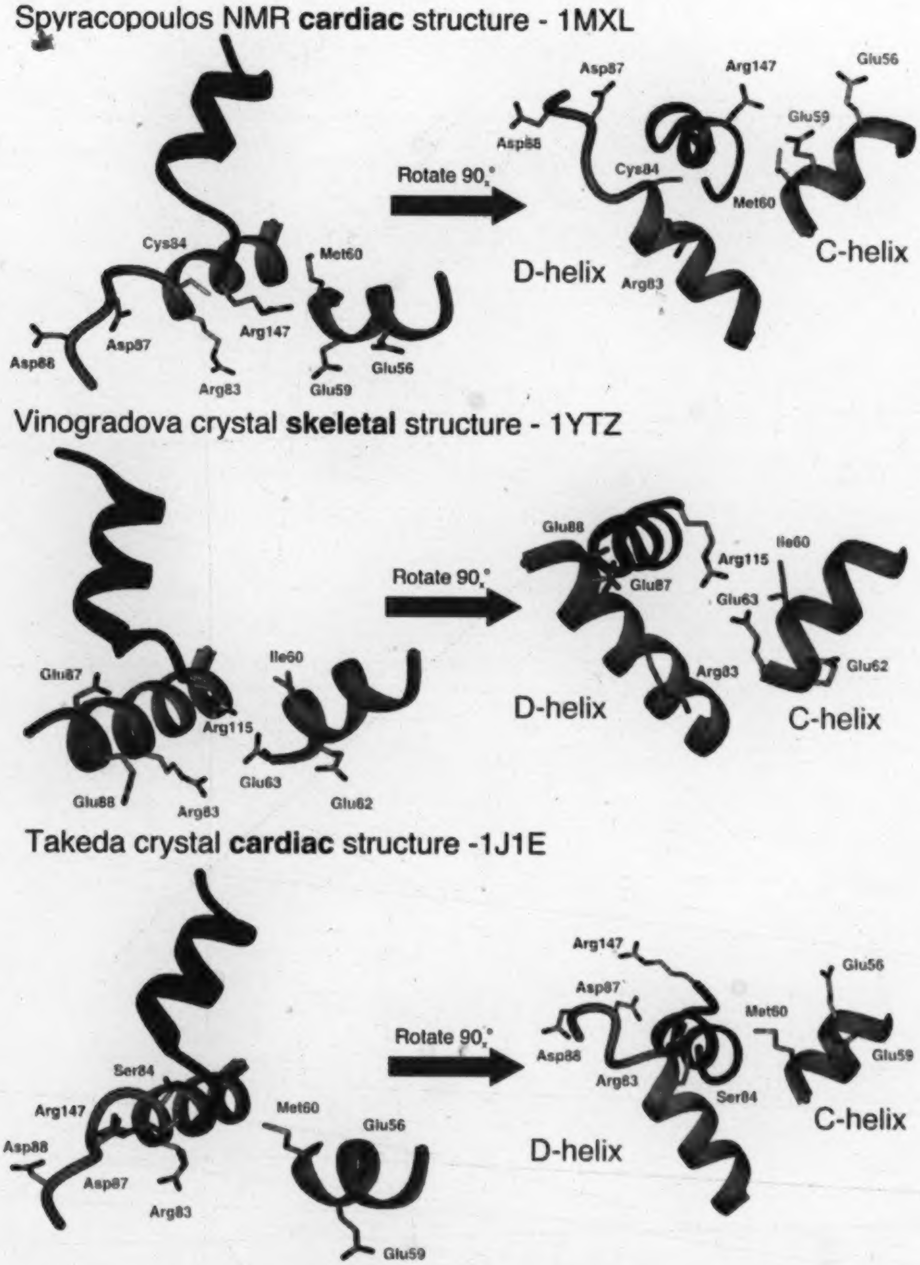
Figure VI-9. **Chemical shift changes of N-lobe upon cSp binding.** (A) Comparison of induced chemical shift changes of cNTnC•Ca^{2+} upon the addition of cSp (maroon), cSp-S149Phos (blue), cSp-R161W (green), and cSp-R161W-S149Phos (red). The chemical shift changes for each residue were calculated by averaging the normalized chemical shift changes of the backbone ^{1}HN and ^{15}N chemical shifts. For a particular nucleus, the normalized chemical shift change of a residue is obtained by dividing the observed shift change by the average shift change for all residues. Thus, ppm = 1 indicates that the chemical shift change for a given residue is equal to the average change for all residues. (B) Chemical shift mapping on the structure of cNTnC in the cNTnC•Ca^{2+}•cSp complex (1MXL.pdb) (*7*): the residues experience major chemical shift changes (ppm ≥ 1.5) are colored red. (C) Residues S149 and R161 are shown on the structure of cNTnC•Ca^{2+}•cSp. The molecular surface of cNTnC•Ca^{2+} (residues 5-84) in the complex is shown with the hydrophobic surface colored yellow, the structure of cSp is shown in the ribbon and specified side chains in the stick representation.

In the high-resolution structure of the cNTnC•Ca^{2+}•cSp complex (*7*), the α-helical segment (residues 151-156) of the 17-residue cSp peptide inserts between the A/B interhelical interface. The N-terminus residues, Arg147-Ser149, interact with the hydrophobic core of cNTnC where Ser149 is located in the center of the core (Figure VI-9C). As such, one negative charge introduced by the phosphorylation of Ser149 (assuming pH=6.7), would be predicted to disrupt the hydrophobic force and lead to weaker binding of cSp for cNTnC•Ca^{2+}. This hypothesis was proven correct via the experimental results recorded in this study. However, it is interesting to note that the affinity of cSp for cNTnC•Ca^{2+} is also weakened by the presence of bepridil in the hydrophobic core although to a smaller extent (~3.5-fold reduction in affinity) (*27*). This indicates that the hydrophobic interaction is the primary anchoring force for cSp binding to cNTnC and stabilizing its opening conformation and it is easily disturbed by the introduction of a negatively charged phosphate group or a partially positively charged agent, such as bepridil. On the other hand, mutation R161W involves a basic to hydrophobic charge change and can be postulated to result an energetically more favorable interaction of the tryptophan side chain with the hydrophobic patch on cNTnC. Unfortunately, in the structure of the cNTnC•Ca^{2+}•cSp complex, the C-terminal residues (Ala160, Arg161, Ala162, and Lys163) of cSp do not interact with cNTnC•Ca^{2+} and extends out of the protein surface (Figure VI-9C). Thus, the mutation R161W may only have minimal effect on the interaction of cSp and cNTnC•Ca^{2+} as reflected in the observed ~1.4-fold affinity decrease. The subtle affinity alteration observed herein is in accordance with a recent surface plasmon resonance-based (SPR) study showing that the affinity of R161W mutant for cTnC is only ~2.6 times that of wild-type cTnI (*21*). Somewhat surprisingly, the effect of S149 phosphorylation is compromised in the presence of R161W mutation, i.e., a ~4-fold affinity decrease by cSp-R161W-S149Phos as compared to a ~10-fold one by cSp-S149Phos. It is possible that the ~10-fold weakening caused by the phosphorylation of Ser149 allows the tryptophan side chain of R161W to rearrange its position slightly such that it makes an energetically more favorable interaction with cNTnC•Ca^{2+} and therefore increasing the apparent affinity.

DISCUSSION

This chapter is two-fold as it is directed towards the N-lobe with regards to both binding of the inhibitory region, and presents an evaluation of the consequences caused by the phosphorylation of Ser149 and FHC mutation R161W on the interaction of cSp with cNTnC. Combined with cIp binding to the C-lobe of cTnC in Chapter IV, and the associated Thr142 phosphorylation and FHC mutation R144G, we have systematically assessed the influences of these alterations on the binding location, stoichiometry, and affinity of these two key regions of cTnI with cTnC.

It is clear based solely upon the {^{1}H, ^{15}N}-HSQC NMR data of the titration of cIp/cSp with the N-lobe that the N-lobe has a preference for cSp over cIp, which results with cSp displacing cIp off the N-lobe binding site. Yet since cIp can illicit structural changes to the N-lobe upon binding, it is clear that residues such as Arg147, Val146 and/or Arg145 may play a role in the binding to the N-lobe. When taken in context with the solved structure of the inhibitory region with that of the C-lobe presented in Chapter III, it was not surprising that the results indicated that the Arg147 played no role in binding to the C-lobe and had altered ^{15}N-NMR relaxation properties, indicating higher flexibility than residues N-terminal to Arg147. The residue Arg147 is therefore a pivotal residue in the cTnI sequence as it is the hypothetical 'division line' between the end of one region (cIp) that binds predominantly via electrostatic interactions, and the start of a region (cSp) that binds predominantly via hydrophobic interactions. It is therefore not surprising that the cSp peptide is highly sensitive to charge placement and modifications that affect the charge density (i.e. R161W and Ser149Phos) as the placement of Ser149 is near the hydrophobic environment of the N-lobe.



Spyracopoulos NMR **cardiac** structure - 1MXL

Cys84
Met60
Asp88
Asp87
Arg147
Arg83
Glu59
Glu56
Rotate 90.°
Asp87
Arg147
Glu56
Asp88
Glu59
Cys84
Met60
D-helix
Arg83
C-helix

Vinogradova crystal **skeletal** structure - 1YTZ

Ile60
Glu87
Arg115
Glu88
Arg83
Glu63
Glu62
Rotate 90.°
Glu88
Arg115
Ile60
Glu87
Glu63
Arg83
Glu62
D-helix
C-helix

Takeda crystal **cardiac** structure -1J1E

Ser84
Arg147
Asp88
Asp87
Arg83
Met60
Glu56
Glu59
Rotate 90.°
Arg147
Asp87
Glu56
Met60
Asp88
Arg83
Ser84
Glu59
D-helix
C-helix

Figure VI-10. **Positioning of Arg147 in the troponin complex**. Ribbon diagrams of the interacting C- and D-helices of the N-lobe (green) of cTnC with Arg147 of the switch region of cTnI (purple). The Spyracopoulos NMR cardiac structure is displayed on the top (*7*), the Vinogradova crystal skeletal structure is displayed in the middle (skeletal numbering – Arg115) (*37*) and the Takeda crystal structure is on the bottom (*36*), with residues in close proximity as indicated.

Close inspection of both the crystal structure and NMR solution structure of the cSp region in complex with the N-lobe reveals a significant difference in the placement of Arg147 of cTnI in relation to the N-lobe of cTnC (Figure VI-10). In the NMR structure, Arg147 is within close proximity to residues Glu56/Glu59/Met60 of the C-helix; in the crystal structure Arg147 is located towards the residues Arg83/Ser84Asp88/Asp89 of the D-helix. The positioning of Arg147 in the NMR solution structure correlates well with my binding results of the inhibitory region as it places Arg147 close within the hydrophobic core of the N-lobe near acidic residues Glu56/Glu59, which may explain the observed affinity and subsequent partial opening for the N-lobe. The crystal structure places Arg147 near the acidic residues Asp87/Asp88 of the D-helix. The placement of Arg147 in this position would not give the results I observed for cIp binding as the only potential contacts are with unstructured residues on the C-terminal end of the D-helix. It is therefore concluded that the positioning of Arg147 in the crystal structure cannot be a physiological preferred conformation, as it does not correlate with the NMR *in vitro* analysis performed independently by both myself and Spyracopoulos (*7*). The recent crystal structure of the *skeletal* troponin complex by Vinogradova reveals a similar placement of Arg115 (skeletal numbering), agreeing with the Spyracopoulos cardiac NMR solution structure, in which Arg115 makes contacts with of the C-helix (i.e. Glu63) (*37*). These results indicate that this interaction is conserved throughout troponin isoforms, and suggests that the Takeda cardiac crystal structure may have improper packing in the crystal lattice, where a non-native positioning of Arg147 is observed.

Residue Ser149 in the cSp region has been recently identified to be the site of p21-activated kinase (PAK), a kinase potentially involved in Ca^{2+}-sensitization of cardiac muscle contraction and cardiac hypertrophy. This chapter has demonstrated that Ser149 phosphorylation reduced the affinity of cSp for cNTnC•Ca^{2+} by ~10-fold. Since Ser149 of cTnI is located in the center of the hydrophobic core of cNTnC•Ca^{2+} (*7*), introduction of a negatively charged phosphate group in this region of cTnI would certainly destabilize its binding to cNTnC•Ca^{2+} and result in reduced affinities. Given that the phosphorylation of Ser149 in cTnI may be involved in a mechanism of Ca^{2+}-sensitization of Triton-skinned cardiac muscle fiber (*17*) and considering that cTnI toggles between cTnC and actin-tropomyosin during the systolic-diastolic cycle, these results suggest a greater reduction in affinity of Ser149 phosphorylated cTnI for actin-tropomyosin as compared to cTnC. Thus cTnI-cTnC interactions would be favored compared to cTnI-actin-tropomyosin interaction resulting in an increase of the Ca^{2+}-affinity for cTnC and thereby the Ca^{2+}-sensitization of the myofilaments. Nonetheless, the dramatic affinity reduction of cTnI for cTnC by the phosphorylation of either Thr142 or Ser149 strongly suggests that kinases such as PKC or PAK modulate cardiac muscle contraction by modifying myofilament protein-protein interaction, contributing to depressed contractility in certain cases and compromised relaxation in other.

In addition to phosphorylation, the cIp-cSp region of cTnI is also prone to missense mutations in diseased heart. Three mutations (R144G, R144Q, and R161W) that co-segregate with FHC are located in this highly conserved region (*19*). Both the R144G and R161W mutations constitute a change in charge and the latter involves the introduction of a large aromatic side chain. The functional abnormalities revealed reduced ability of R144G and R161W mutants to inhibit actomyosin ATPase and an increase in Ca^{2+} sensitivity of actomyosin ATPase regulation (*21-25*). If these functional differences manifest themselves *in vivo* there will be impairment of relaxation of cardiac muscle, and this altered contractility may provide a hypertrophic stimulus eventually leading to cardiac dysfunction. This is supported by a mouse model study that shows that at the whole organ level, contractile function is enhanced but relaxation is compromised in mouse harboring the R144G mutation (*22*). Because cTnI toggles between cTnC and actin-tropomyosin, it seems likely that an important mechanism of cTnI-induced cardiac dysfunction by the R144G and R161W mutations lies in altered cTnI interactions with cTnC, actin-tropomyosin or both. The binding of cTnI to actin-tropomyosin inhibits actomyosin ATPase activity, depresses cross-bridge cycling, and prevents contraction. The inhibition is released as cTnI binding shifts from actin-tropomyosin to cTnC, an action favored by an increase in cytosolic Ca^{2+} concentration during systole, when Ca^{2+} binds to the N-lobe of cTnC. Thus, it seems that the R144G or R161W mutated cTnI either has reduced affinity for actin-tropomyosin or enhanced affinity for cTnC. The net effect would be the diminished inhibition of actomyosin ATPase activity. Interestingly, our present results indicate that the affinity of cIp for cCTnC•2Ca^{2+} is reduced by ~6-fold by the mutation R144G and that of cSp for cNTnC•Ca^{2+} is reduced ~1.4-fold by the mutation R161W. In light of the above argument, these results suggest a greater reduction in affinity of R144G and R161W mutants for actin-tropomyosin compared to cTnC, such that cTnI-cTnC interaction is favored over cTnI-actin-tropomyosin interaction, similar as in the case of Ser149 phosphorylation.

The fact that the cIp and cSp segments of cTnI are subject to modifications by both phosphorylation and mutation prompted me to examine the binding of cIp-R144G-T142Phos to cCTnC•2Ca^{2+} (Chapter IV) and that of cSp-R161W-S149Phos to cNTnC•Ca^{2+}. The results show that T142Phos and R144G modifications have a cumulative effect on reducing the affinity of cIp for cCTnC•2Ca^{2+}; however, the destabilizing factor introduced by Ser149 phosphorylation on the interaction of cSp and cNTnC•Ca^{2+} is compromised somehow by the R161W mutation. Thus, it seems that the FHC mutations can either enhance or suppress the effect of phosphorylation depending on the specificity of the interactions. So far there is limited functional data available from literature linking the FHC mutations with cTnI phosphorylation. Current data will be helpful for future work in understanding the mechanism of cTnI function involving both FHC mutations and PKC or PAK phosphorylation.

In summary, cTnI peptides have been used to examine the consequences of cTnI modification on its interaction with cTnC in Chapters IV & V. The results indicate that a given modification of cTnI can have a pronounced effect on the binding affinity of cTnI for cTnC. These perturbations in protein affinity may correlate with the impairment of cTnI function in heart muscle contraction. Clearly, further studies by employing the intact cTnI protein, the entire troponin complex, and even the entire thin filament are needed to generate a more complete picture in the understanding of the mechanism by which myofilament protein modifications promote cardiac dysfunction.

REFERENCES

1. Head, J. F., and Perry, S. V. (1974) *Biochemical Journal 137*, 145-154.
2. Talbot, J. A., and Hodges, R. S. (1979) *J. Biol. Chem. 254*, 3720-3723.
3. Sia, S. K., Li, M. X., Spyracopoulos, L., Gagné, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) *J. Biol. Chem. 272*, 18216-18221.
4. Mercier, P., Ferguson, R. E., Irving, M., Corrie, J. E., Trentham, D. R., and Sykes, B. D. (2003) *Biochemistry 42*, 4333-48.
5. McKay, R. T., Tripet, B. P., Hodges, R. S., and Sykes, B. D. (1997) *J. Biol. Chem. 272*, 28494-28500.
6. Gagné, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) *Nature Struct. Biol. 2*, 784-789.
7. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.
8. Farah, C. S., Miyamoto, C. A., Ramos, C. H. I., da Silva, A. C. R., Quaggio, R. B., Fujimori, K., Smillie, L. B., and Reinach, F. C. (1994) *Journal of Biological Chemistry 269*, 5230-5240.
9. Brown, L. J., Sale, K. L., Hills, R., Rouviere, C., Song, L., Zhang, X., and Fajer, P. G. (2002) *Proc Natl Acad Sci U S A 99*, 12765-70.
10. Spyracopoulos, L., Li, M. X., Sia, S. K., Gagné, S. M., Chandra, M., Solaro, R. J., and Sykes, B. D. (1997) *Biochemistry 36*, 12138-12146.
11. Vassylyev, D. G., Takeda, S., Wakatsuki, S., Maeda, K., and Maeda, Y. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 4847-4852.
12. Ferguson, R. E., Sun, Y. B., Mercier, P., Brack, A. S., Sykes, B. D., Corrie, J. E., Trentham, D. R., and Irving, M. (2003) *Mol Cell 11*, 865-74.
13. Gagné, S. M., Tsuda, S., Spyracopoulos, L., Kay, L. E., and Sykes, B. D. (1998) *J. Mol. Biol. 278*, 667-686.
14. Heller, W. T., Abusamhadneh, E., Finley, N., Rosevear, P. R., and Trewhella, J. (2002) *Biochemistry 41*, 15654-63.
15. Heller, W. T., Finley, N. L., Dong, W. J., Timmins, P., Cheung, H. C., Rosevear, P. R., and Trewhella, J. (2003) *Biochemistry 42*, 7790-7800.
16. Solaro, R. J., and Rarick, H. M. (1998) *Circ Res 83*, 471-80.
17. Buscemi, N., Foster, D. B., Neverova, I., and Van Eyk, J. E. (2002) *Circ Res 91*, 509-16.
18. Redwood, C. S., Moolman-Smook, J. C., and Watkins, H. (1999) *Cardiovasc Res 44*, 20-36.
19. Hernandez, O. M., Housmans, P. R., and Potter, J. D. (2001) *J Appl Physiol 90*, 1125-36.
20. Murphy, A. M., Kogler, H., Georgakopoulos, D., McDonough, J. L., Kass, D. A., Van Eyk, J. E., and Marban, E. (2000) *Science 287*, 488-91.
21. Elliott, K., Watkins, H., and Redwood, C. S. (2000) *J Biol Chem 275*, 22069-74.
22. James, J., Zhang, Y., Osinska, H., Sanbe, A., Klevitsky, R., Hewett, T. E., and Robbins, J. (2000) *Circ Res 87*, 805-11.

23. Takahashi-Yanaga, F., Morimoto, S., Harada, K., Minakami, R., Shiraishi, F., Ohta, M., Lu, Q. W., Sasaguri, T., and Ohtsuki, I. (2001) *J Mol Cell Cardiol 33*, 2095-107.
24. Deng, Y., Schmidtmann, A., Redlich, A., Westerdorf, B., Jaquet, K., and Thieleczek, R. (2001) *Biochemistry 40*, 14593-602.
25. Lang, R., Gomes, A. V., Zhao, J., Housmans, P. R., Miller, T., and Potter, J. D. (2002) *J Biol Chem 277*, 11670-8.
26. Burkart, E. M., Sumandea, M. P., Kobayashi, T., Nili, M., Martin, A. F., Homsher, E., and Solaro, R. J. (2003) *J Biol Chem 278*, 11265-72.
27. Wang, X., Li, M. X., and Sykes, B. D. (2002) *J Biol Chem 277*, 31124-33.
28. Gagné, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) *Protein Science 3*, 1961-1974.
29. Li, M. X., Gagné, S. M., Tsuda, S., Kay, C. M., Smillie, L. B., and Sykes, B. D. (1995) *Biochemistry 34*, 8330-8340.
30. Golosinska, K., Pearlstone, J. R., Borgford, T., Oikawa, K., Kay, C. M., Carpenter, M. R., and Smillie, L. B. (1991) *J. Biol. Chem. 266*, 15797-15809.
31. Tripet, B. P., Van Eyk, J. E., and Hodges, R. S. (1997) *J. Mol. Biol. 271*, 728-750.
32. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.
33. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.
34. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
35. Johnson, A. J., and Blevins, R. A. (1994) *Journal of Biomolecular NMR 4*, 603-614.
36. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
37. Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) *Proc Natl Acad Sci U S A.*

CHAPTER VII

Troponin I interactions within a 38kDa Tn complex

OVERVIEW

This chapter extends the NMR structural studies of troponin I that have been reported in previous chapters, which relied exclusively on peptides and smaller domains of the troponin complex in an attempt to limit molecular weight and/or maximize solubility for NMR experiments. This chapter expands these studies and concentrates on a 38kDa troponin complex with the intent of relating it to the crystal structure. The location and flexibility of the inhibitory region in relation to the rest of the complex is a key interest in determining if the *in vitro* solution troponin complex matches the Takeda crystal structure (1J1E.pdb) and/or the Lindhout model. Inference for biological function and physiological effects can then be addressed in an attempt to understand the role of phosphorylation as a regulatory control, and altered thin filament interactions in the presence of a troponin I mutation(s) found in familial hypertrophic cardiomyopathy.

The results reported in this chapter are unfinished, unpublished, and are thus reported as a 'work-in-progress'. Previously reported methodologies for the production of a 'core' troponin complex have been modified in an attempt to optimize a sample for NMR studies and have a correctly folded complex. The availability of an ultra-high frequency 800MHz NMR spectrometers, coupled with newer pulse sequence development such as transverse relaxation optimized spectroscopy (TROSY), have made possible NMR studies of high molecular weight complexes that were unattainable only 10 years ago. Working with $^{13}C/^{15}N$-cTnC and/or $^{13}C/^{15}N/^{2}H$-cTnI, I have been successful in producing the first NMR spectra of a 38kDa troponin complex consisting of cTnC•cTnI$_{34\text{-}163}$•cTnT$_{226\text{-}271}$. Spectral assignments of this complex are currently incomplete; however, inferences to secondary structure and molecular architecture can be made. This chapter reports preliminary data on a project that will require further years of study.

INTRODUCTION

Within this thesis and the literature, tremendous amounts of structural work have been reported of domain(s) within the troponin complex (*1-29*); however, structural studies of the individual domains can only yield so much information before the pursuit of larger, more complex systems becomes a necessity. The start of the 21 century has witnessed a sharp rise in both the quality and quantity of structural studies of the troponin complex due to the advent of more sophisticated technologies and novel ideas. The past two years (2003-2005) have produced a number of high-resolution Small Angle Neutron Scatting (SANS) and x-ray crystal structures, which has been key in understanding the components of the muscle molecular motor. In 2003, Soichi Takeda and associates produced the first quality crystal structure of a 52kDa cardiac troponin complex, which revealed structural insights into protein fold and domain association (*30*). In 2003, William Heller and associates established low-resolution domain orientations of the cardiac troponin complex using SANS techniques (*16, 17*). In 2004, Fumiko Matsumoto and associates provided *in situ* calcium-induced conformational changes of troponin C in native thin filaments by SANS (*31*). In 2005, William King and associates produced the solution structure of the chicken skeletal muscle troponin complex via a combination of SANS and x-ray scattering (*32*). In 2005, Maia Vinogradova and associates produced a skeletal troponin complex by crystallography, which was the equivalent of the Takeda cardiac structure (*33*).

The structural insights obtained from crystallography have provided a framework for dissecting protein-protein interactions on the thin filament during cardiac contraction. Unfortunately, these crystal structures leave many questions unanswered. The underlying difficulty in pursuing crystallography of an intrinsically flexible protein complex such as the troponin complex lies in the fact that the crystal structure may only 'freeze' a single conformation. This single conformation may not fully explain the movements and flexibilities required for calcium signaling during a contraction cycle. The authors of these crystal structures offer models of domain movement(s); however, there was no way of obtaining the dynamics, a necessary component in the understanding of cardiac contraction.

The recent advances in technology also applies to the field of NMR, as novel pulse sequence technology (i.e. TROSY) that takes advantage of increased magnetic field strengths (i.e. 800MHz), makes it possible to investigate high molecular weight protein complexes such as the troponin complex (*34, 35*). NMR is a perfect complementary technique to the crystal structures as it is uniquely able to report ^{15}N-backbone and/or ^{13}C-sidechain dynamics via the exploitation of magnetic relaxation. It is through this technique that I planned on investigating a cardiac troponin complex of approximately 40kDa consisting of all three troponin subunits (I, T & C).

Previous chapters in this thesis have indicated the importance of troponin I in the cardiac contraction/relaxation cycles. Although past cTnI studies have focused predominantly on the troponin C subunit and structurally induced effects (*9, 10, 15*), this chapter will provide the first reported investigation of full-length troponin I in the troponin complex by NMR. The crystal structure was investigated to limit the flexibility of domains in attempts to make a well-behaved solution sample for NMR. Regions C-terminal to the switch region (residues 164-210) and N-terminal to the cRP40 region (residue 1-33) of cTnI were omitted as they do not make direct contact with the complex, leaving a functional domain of 129 residues (cTnI$_{34-163}$) in the troponin complex. Protocols for the cloning, expression and purification of the cTnI$_{34-163}$ and the cTnT$_{197-298}$ subunits, and the subsequent incorporation of cTnI$_{34-163}$ into a high molecular weight troponin complex (i.e. 38.8kDa and/or 45.4kDa) are detailed in the experimental protocols.

EXPERIMENTAL PROCEDURES

VII-A. Cloning, expression and purification of $cTnT_{197-298}$

cDNA encoding full length cTnT gene in a pET3a vector was obtained. The DNA encoding only the region $cTnT_{197-298}$ was amplified using *Invitrogen*-Plantinum Blue PCR SuperMix (94°C for 30s, 52°C for 30s, 72°C for 60s, repeat 30X). A forward primer (F1) was designed as 5'-CCCCCGCATATGATCCAGAAGCAGGCC-3' with a 5'-Nde1 restriction cut site and a reverse primer (R1) was designed as 5'-GAAGGGATCCTATTTCCAGCGCCCGGTGAC-3' with a 3'-BamHI restriction cut site. Ligation, transformation, selection and verification of positive plasmids directly follows the protocol for the GEV1 expression vector outlined in Chapter II, except that PCR products were ligated into un-altered pET3a-BamHI/NdeI cut vector and identification of positive colonies was confirmed using the restriction enzyme NsiI. The expression and induction of the $cTnT_{197-198}$ protein mirrors the protocol outlined in Chapter III of this thesis for cCTnC expression.

Following centrifugation of the cell culture, the pellet was frozen at -20°C overnight in 50mM Tris buffer (pH 8.0), thawed to room temperature, and resuspended. Solid urea was added to the cellular pellet solution to a final concentration of 6M, along with 15mg of DTT, 420mg of citrate, broken in a French pressure cell, and centrifuged at 17500rpm for 3hrs and passed through a 0.22*u*m filter to remove all insoluble cellular debris. Supernatant was loaded onto a fast-flow CM-sepharose exchange column equilibrated with CM-buffer-A (20mM citric acid, 6M urea, 2mM EDTA, 1mM DTT, pH 8.0) and the column was thoroughly rinsed with CM-buffer-A. An increasing salt-gradient was applied using 500mL each of CM-buffer-A and CM-buffer-B (20mM citric acid, 6M urea, 2mM EDTA, 1mM DTT, 0.6M NaCl, pH 8.0), and fraction elution of protein was monitored by SDS-PAGE gel electrophoresis. Fractions that contained the presence of $cTnT_{197-298}$ protein (based upon SDS-PAGE gel analysis) were pooled together and lyophilized to dryness.

Solid, cTnT$_{197-298}$ powder was further purified using a Superdex 30 column. Protein powder was dissolved in 2mL of Superdex column buffer (20mM citric acid, 6M urea, 2mM EDTA, 1mM DTT, 0.3M NaCl, pH 8.0), loaded onto a Superdex 30 column using gravity flow. All elution fractions were monitored using SDS-PAGE gel electrophoresis, and fractions containing pure cTnT$_{197-298}$ were pooled. cTnT$_{197-298}$ was de-salted using a G-25 Sephadex de-salt column equilibrated with 10mM NH_4HCO_3 (pH 8.0), and was lyophilized to dryness.

VII-B. Cloning, expression and purification of cTnI$_{34-163}$

cDNA encoding full length cTnI gene (C80S, C97S) in a pET3a vector was obtained as a kind gift from the Dr. Larry Smillie laboratory (Department of Biochemistry, University of Alberta). The DNA encoding only the region cTnI$_{34-163}$ was amplified using *Invitrogen*-Plantinum Blue PCR SuperMix (94°C for 30s, 55°C for 30s, 72°C for 60s, repeat 30X). A forward primer (F2) was designed as 5'-CCCCCGCATATGGCCAAGAAAAAATCTAAG-3' with a 5'-Nde1 restriction cut site and a reverse primer (R2) was designed as 5'-GAAGGGATCCTAAGCCCGGGCCCCCAGCAG-3' with a 3'-BamHI restriction cut site. Ligation, transformation, selection and verification of positive plasmids directly follows the protocol for the GEV1 expression vector outlined in Chapter II, except that PCR products were ligated into un-altered pET3a-BamHI/NdeI cut vector and identification of positive colonies was confirmed using the restriction enzyme XhoI. The expression and induction of the cTnT$_{197-198}$ protein mirrors the protocol outlined in Chapter III of this thesis for cCTnC expression. The expression and induction of the cTnI$_{34-163}$ protein mirrors the protocol outlined in Chapter III of this thesis for cCTnC expression. This expression protocol results in cTnI$_{34-163}$ sequestration within bacterial inclusion bodies.

Following centrifugation of the cell culture, the pellet was re-suspended in Lysis buffer (50mM Tris, 150mM NaCl, 5mM EDTA, 1mM DTT, 0.5mM PMSF, pH 7.5), broken in a French pressure cell and centrifuged at 10000rpm for 15min. The supernatant was decanted, and the pellet was re-dispersed by homogenization in Re-dispersion buffer (1M urea, 0.5% Triton X-100), followed by centrifugation at 10000rpm for 15 minutes, and repeated 4 times. The pellet was then solubilized in Equilibration buffer-1 (50mM Tris, 8M urea, 10mM NaCl, 1mM EDTA, 1mM DTT, pH 8.0) and centrifuged at 18000rpm for 1hr. The supernatant was then loaded onto a CM-sepharose column equilibrated with Equilibration buffer-1. An increasing salt-gradient (1L) was established for elution of cTnI$_{34-163}$, using 500mL of Equilibration buffer-1 and 500mL of Equilibration buffer-2 (500mM NaCl, 50mM Tris, 8M urea, 10mM NaCl, 1mM EDTA, 1mM DTT, pH 8.0). Column fractions were monitored by SDS-PAGE and fractions containing pure cTnI$_{34-163}$ were pooled and dialyzed against 1% formic acid for 24hrs at 4°C, and lyophilized to dryness.

For incorporation of isotopic labels (i.e. ^{13}C, ^{15}N and/or ^{2}H) during expression of the protein, the protocol outlined for cCTnC expression in Chapter III was followed, where M9 minimal media (50mM Na_2HPO_4, 25mM KH_2PO_4, 10mM NaCl, pH 7.3) supplemented with ^{13}C-glucose, ^{15}N-$(NH_4)_2SO_4$ and/or D_2O was used.

VII-C. Sequencing cTnT$_{197-298}$ and cTnI$_{34-163}$ pET3a vectors

All ligated vectors confirmed positive by restriction digest were subjected to further confirmation by DNA sequencing using the T7-terminator primer. The sequence of the T7-terminator primer used was 5'-TATGCTAGTTATTGCTCAG-3'. Sequencing results are listed in *Appendix A-3*.

VII-D. Synthesis of cTnT$_{236-271}$ peptide

The synthesis of peptide cTnT$_{226-271}$ is as previously described as for wild-type cIp in Chapter III of this thesis (*1*, *6*). cTnT$_{226-271}$ is a 46 residue peptide with a sequence; acetyl-`LNEDQLREKAKELWQSIYNLEAEKFDLQEKFKQQKYEINVLRNRI`-amide. After synthesis and purification, this peptide was verified by MALDI-TOF mass spectrometry and lyophilized twice to remove all traces of organic solvents.

VII-E. Production of a high molecular weight 45.4kDa troponin complex

The protocol for making a high molecular weight complex is modified from that reported in Heller (*17*). Purified, lyophilized cTnC, $cTnI_{34\text{-}163}$ and $cTnT_{197\text{-}298}$ were each weighed separated as to create a 0.3mM sample in 1.0mL of buffer at molar concentrations of 1•1•1.1. Samples were created using ^{13}C, ^{15}N, and/or ^{2}H isotope incorporation in either cTnC and/or cTnI. Weighed samples were combined and initially dissolved in 3mL of Urea buffer (8M urea, 200mM NaCl, 25mM Tris, 1mM EDTA, 1mM DTT, 10mM $CaCl_2$, 0.01% NaN_3, pH 7.5) with small amounts of 0.1M HCl and 10% formic acid added until all components completely dissolved. The sample solution was then transferred into dialysis tubing (MWCO=1000) and sealed, ensuring no leaks. The sample was then placed in 500mL of Urea buffer at 4°C with stirring and a gradual gradient against 5L of Dialysis buffer A (350mM KCl, 25mM Tris, 1mM EDTA, 1mM DTT, 10mM $CaCl_2$, 0.01% NaN_3, pH 7.5) was established at 104mL/hr for a total of 48hrs, while maintaining a constant volume of 500mL. To reduce the total salt concentration, the sample was then dialyzed against 5L of Dialysis buffer B (100mM KCl, 25mM Tris, 1mM EDTA, 1mM DTT, 10mM $CaCl_2$, 0.01% NaN_3, pH 6.7) using a gradient of 104mL/hr for 48hrs while maintaining a constant volume of 500mL. A small amount of white precipitate was observed once the KCl concentration was reduced below 200mM, and a large amount of precipitate occurred once the concentration went below 150mM. All future production protocols were adjusted for Dialysis buffer B to contain 150mM KCl instead of 100mM KCl to avoid precipitation issues of the complex. Following dialysis, the sample volume was reduced from 3mL to 1mL using 1.5mL eppendorf spin-filter tubes equipped a 1000 MWCO membrane. Samples were centrifuged at 2Krpm for 1hr at room temperature.

VII-F. Verification of a 45.4kDa troponin complex by electrophoresis

To remove small quantities of unbound troponin subunits, the sample was purified using a Superdex-200 column (1.5m length, 0.9mm diameter) using gravity flow, equilibrated with Dialysis buffer B. The elution rate was measured as approximately

0.5mL/hr and all elution fractions were checked with both non-denaturing (native) and denaturing SDS-PAGE. Non-denaturing 8% PAGE gels were created using 1X separating Gel buffer (25mM Tris, 80mM Glycine, 5mM $CaCl_2$, 10% w/w Glycerol, pH 8.6) and run using 1X running buffer (25mM Tris, 80mM Glycine, 5mM $CaCl_2$). The positive fractions containing the 45.4kDa complex were pooled together and sample volume was decreased to 550*u*L using 1.5mL eppendorf spin-filter tubes equipped a 1000 MWCO membrane. 100% D_2O (50*u*L) and DSS (10*u*L of 10mM) were added to the sample and 500*u*L was added directly to a clean, 5mm NMR tube.

VII-G. Production of a 38.8kDa troponin complex

Production of a 38.8kDa complex was created using identical procedures outline for the 45.4kDa complex, however the synthetic peptide $cTnT_{226\text{-}271}$ was used in place of $cTnT_{197\text{-}298}$. Purified, lyophilized cTnC, $cTnI_{34\text{-}163}$ and $cTnT_{226\text{-}271}$ were each weighed separated as to create molar ratios of 1•1•1.1. Samples were created using ^{13}C, ^{15}N, and/or ^{2}H isotope incorporation in either cTnC and/or cTnI.

VII-H. NMR spectroscopy

All NMR data used in this study were acquired at 30°C using Varian INOVA 500 MHz, Unity 600 MHz, and INOVA 800 MHz spectrometers. All three spectrometers are equipped with triple resonance probes and Z-axis pulsed field gradients (XYZ gradients for the 800 MHz). Experiments performed are a combination of Lewis Kay style and BioPack pulse sequences. All NMR experiments run for all samples are detailed in *Table VII-1*. Spectral processing of all data was performed using the software package NMRPipe by Frank Delaglio (*36*), and spectral viewing, assignments and relaxation fit using the software package nmrView by Bruce Johnson (*37*). All samples were referenced directly (indirectly for $^{13}C/^{15}N$ dimensions) to a 0.2mM internal standard of 2,2-Dimethyl-2-silapentane-5-sulfonate (Cambridge Isotope Laboratories).

VII-I. ^{15}N-NMR T1 and T2 relaxation parameters

All relaxation data were acquired using an INOVA 500MHz NMR spectrometer at 30°C. The delay times used for ^{15}N-T_2 relaxation measurements were set to 16.61 × *n* ms (where $n = 1, \ldots, 11$). All other parameters were set equal to the {^{1}H, ^{15}N}-HSQC parameters previously in *Table VII-I*.

Table VII-I

NMR spectra acquired and experimental conditions

A. 45.4kDa sample ^{13}C/^{15}N-cCTnC•cTnI$_{34\text{-}163}$•cTnT$_{197\text{-}298}$

Experiment Name	Nuclei[a]	^{1}H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D – {^{1}H, ^{15}N}-HSQC	^{1}H,^{15}N	500	32	878	128		7000	1800			*(38)*

B. 38.8kDa sample ^{15}N-cCTnC•cTnI$_{34\text{-}163}$•cTnT$_{226\text{-}271}$

Experiment Name	Nuclei[a]	^{1}H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D – {^{1}H, ^{15}N}-HSQC	^{1}H,^{15}N	500/800	96/32	878/1166	128/128		6000/10400	1800/3242			*(38)*
2D – {^{1}H, ^{15}N}-TROSYHSQC	^{1}H,^{15}N	800	32	6168	128		10400	3242			*(35)*

C. 45.4kDa sample ^{15}N-cTnI$_{34\text{-}163}$•cTnC•cTnT$_{197\text{-}298}$

Experiment Name	Nuclei[a]	^{1}H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D – {^{1}H, ^{15}N}-HSQC	^{1}H,^{15}N	600	32	768	128		8000	2000			*(38)*
2D – {^{1}H, ^{15}N}-TROSYHSQC	^{1}H,^{15}N	600	32	3072	128		8000	2000			*(35)*
3D - ^{15}N-edited NOESY-HSQC	^{1}H,^{1}H,^{15}N	600	8	1024	96	32	7800	7800	2000	150	*(39)*

D. 38.8kDa sample ^{15}N/^{2}H-cTnI$_{34\text{-}163}$•cTnC•cTnT$_{197\text{-}298}$

Experiment Name	Nuclei[a]	^{1}H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D – {^{1}H, ^{15}N}-HSQC	^{1}H,^{15}N	600/800	32/32	768/1076	128/192		6000/9600	1800/3242			*(38)*
2D – {^{1}H, ^{15}N}-TROSYHSQC	^{1}H,^{15}N	600/800	96/32	3072/4304	128/128		7800/9600	2000/3242			*(35)*
2D – ^{15}N-T$_2$	^{1}H,^{15}N	600	32	768	128		6000	1800			*(40)*

E. 38.8kDa sample ^{13}C/^{15}N-cTnI$_{34\text{-}163}$•cTnC•cTnT$_{197\text{-}298}$

Experiment Name	Nuclei[a]	^{1}H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D – {^{1}H, ^{15}N}-HSQC	^{1}H,^{15}N	600/800	32/32	768/1076	128/192		8000/9600	2000/3242			*(38)*
2D – {^{1}H, ^{15}N}-TROSYHSQC	^{1}H,^{15}N	600	32	3072	128		8000	2000			*(35)*
3D - ^{15}N-edited NOESY-HSQC	^{1}H,^{1}H,^{15}N	800	24	1344	96	26	12001	8500	2100	100	*(39)*
3D - ^{15}N-edited DIPSI-HSQC	^{1}H,^{1}H,^{15}N	500	16	768	72	17	6000	6000	1250	80.2	*(39)*
3D – HNCO	^{1}H,^{13}C,^{15}N	600	8	844	128	30	6600	1500	1216		*(41)*
3D – CBCA(CO)NNH	^{1}H,^{13}C,^{15}N	600	24	1024	50	32	8398	12069	2432		*(42)*
3D – HNCACB	^{1}H,^{13}C,^{15}N	600	16	1024	74	32	8398	12069	2432		*(42)*
2D – ^{15}N-T$_2$	^{1}H,^{15}N	500	16	768	128		6000	1800			*(40)*

F. 38.8kDa sample ^{13}C/^{15}N/^{2}H-cTnI$_{34\text{-}163}$•cTnC•cTnT$_{197\text{-}298}$

Experiment Name	Nuclei[a]	^{1}H freq[b]	nt[c]	x-pts[d]	y-pts	z-pts	x-sw	y-sw	z-sw	Mix[e]	Ref.
2D – {^{1}H, ^{15}N}-HSQC	^{1}H,^{15}N	600/800	128/16	922/1344	128/96		7200/11999	1650/2500			*(38)*
2D – {^{1}H, ^{15}N}-TROSYHSQC	^{1}H,^{15}N	600/800	32/16	3688/5376	128/96		7200/11999	1650/2500			*(35)*
3D – HNCA	^{1}H,^{13}C,^{15}N	800	16	986	64	32	8800	6034	2500		*(41)*
3D – HNCO	^{1}H,^{13}C,^{15}N	500	16	730	50	28	5000	1500	1216		*(41)*
2D – ^{15}N-T$_2$	^{1}H,^{15}N	600	32	896	128		7200	1250			*(40)*

[a] The nucleus acquired in each dimension (i.e. ^{1}H,^{15}N indicates ^{1}H in x-dimension and ^{15}N in y-dimension).

[b] The frequency of ^{1}H in MHz.

[c] the number of transients acquired for each FID

[d] x,y,z-pts and sw is the number of complex points and sweep width in each respective dimension (x is the directly detected dimension)

[e] mixing times are given in milliseconds. In the case of DISPI experiments, this is the spin-locking time.

RESULTS

Studies of the troponin complex were initially pursued using $^{15}N/^{13}C$-cTnC, as it is expressed in relative high yields (80mg/L of minimal media), easily purified (see Chapter V) and provides superior $^{1}H/^{15}N$-NMR spectral dispersion for initial characterization(s). Once a stable protocol for construction of a core troponin complex was established, focus was shifted to ^{15}N-, ^{13}C- and/or ^{2}H-$cTnI_{34-163}$, which expressed at lower levels, required a broader labor/time-intensive purification, and possesses poorer spectral dispersion. As well, due to solubility concerns, samples were kept at a concentration of 300*u*M, requiring a higher number of scans and higher NMR fields (i.e. 800MHz) to maximize signal to noise and produce higher quality line-shape.

VII-A. Choice of peptide $cTnT_{226-271}$ for use in the troponin complex

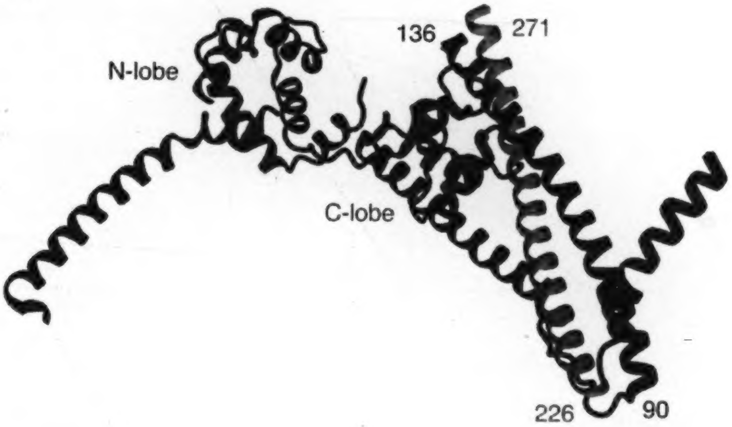


Figure VII-1. **Coiled-coil interaction of cTnT and cTnI subunits**. The crystal structure (1J1E.pdb) is shown as a ribbon diagram with the troponin complex drawn in red, and the coiled-coil domains drawn in green (cTnT) and blue (cTnI).

The purified 12.2kDa $cTnT_{197-298}$ recombinant protein was poorly behaved in solution and displayed evidence of precipitation problems in urea buffers. The addition of organic solvents (i.e. TFE-D_3, CD_3OH, DMSO-D_6) made no appreciable increase in the solvation properties of this domain. Investigation of the Takeda crystal structure

(1J1E.PDB) reveals that only residues cTnT$_{226\text{-}271}$ are important in making a coiled-coil motif with cTnI$_{90\text{-}136}$ in the larger complex (see Figure VII-1).

The synthetic peptide proved to have superior solubility when compared to the recombinant cTnT domain. I therefore decided to make use of the 5.6kDa cTnT$_{226\text{-}271}$ peptide instead of recombinant cTnT$_{197\text{-}298}$ to ease the production of a high-molecular weight complex due to solubility concerns, and to lower the total molecular weight from 45.4kDa to 38.8kDa. Verification of proper complex formation using non-denaturing PAGE analysis displayed a large band at approximately 40kDa (corresponding to the 38.8kDa complex), indicating a 1•1•1 complex formation. Contaminating uncomplexed protein bands corresponding to 33.1kDa (~~cTnC~~•cTnI$_{34\text{-}163}$) and 18.4kDa (unbound cTnC) were observed along with the 38.8kDa ternary complex (Figure VII-2, Lane 3). Further purification by superdex gel filtration produced a pure ternary complex (Figure VII-2, Lane 4). Native PAGE gels were observed to run at a direct proportionality of distance migration to protein molecular weight. Investigation of troponin complexes with incorporation of either cTnT$_{226\text{-}271}$ or cTnT$_{197\text{-}298}$ revealed {^{1}H, ^{15}N}-TROSYHSQC spectra identical to that observed in Figure VII-3C.

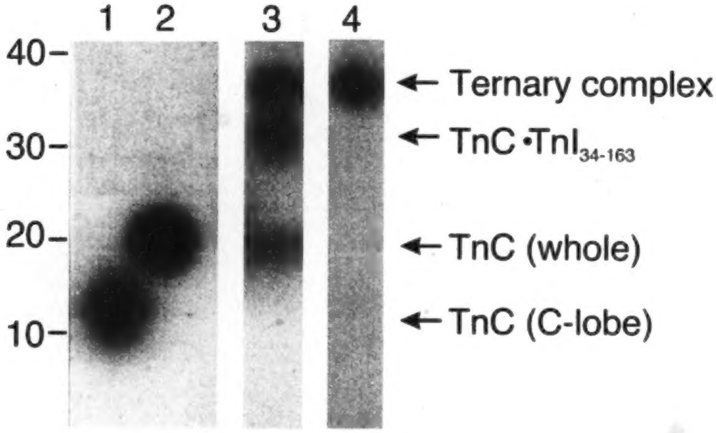


Figure VII-2. **Native PAGE gel strip plots of a 38.8kDa ternary troponin complex**. Lanes 1 and 2 display molecular weight markers TnC (whole and C-lobe), Lane 3 reveals contaminating protein components after ternary complex formation. Lane 4 demonstrates ternary complex purification by superdex gel filtration.

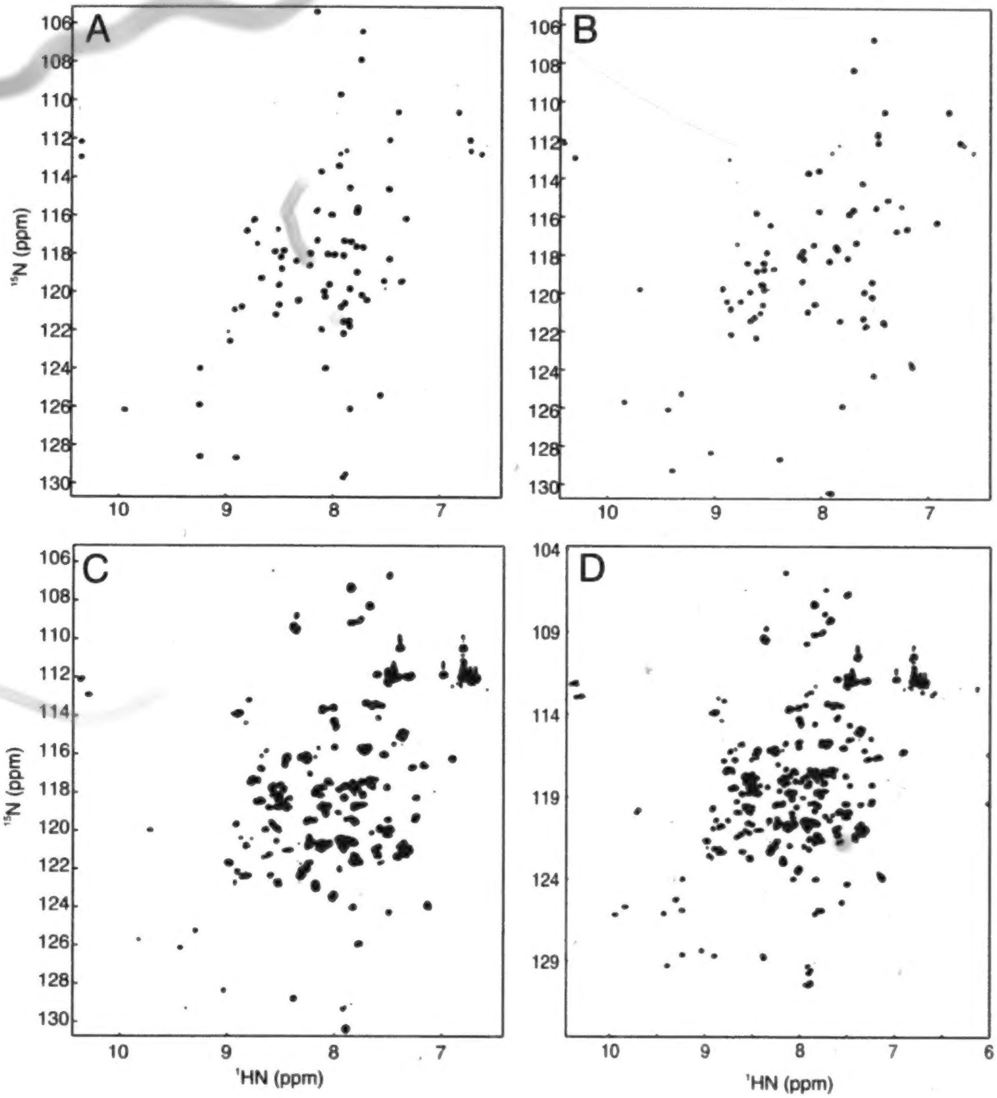


Figure VII-3. **{^{1}H, ^{15}N}-HSQC of ^{15}N-cTnC binding to troponin**. The progressive amide shifts are shown for (A) C-lobe @ 500MHz (black – unbound), (B) C-lobe @ 500MHz bound to cRP40 (red), (C) cTnC @ 800MHz bound to cTnI$_{34-163}$ and cTnT$_{226-271}$ (TROSY – blue) and (D) an overlay of the three spectra.

Figure VII-3 shows the $\{^1H, ^{15}N\}$-TROSYHSQC spectra of the complex, which reveals significant chemical shift perturbations of the backbone amide nuclei when compared to cCTnC alone (see Chapter III) and cCTnC bound to only the cRP40 region (see Chapter V). This can be seen clearly by looking at an expanded region of these spectra corresponding to Gly110 and Gly146 of cTnC (Figure VII-4). It was not surprising to see a change in the positioning of the two Gly residues from cRP40 bound to that of the complete troponin complex, as the coiled-coil region makes close contacts to the calcium binding loops (Sites III & IV). The broader lines seen for the complex are indicative of a higher molecular weight. Based upon these observations and supplemental molecular weight verification by non-denaturing PAGE analysis, it was concluded that a reproducible protocol for the production of a 38.8kDa complex had been established.

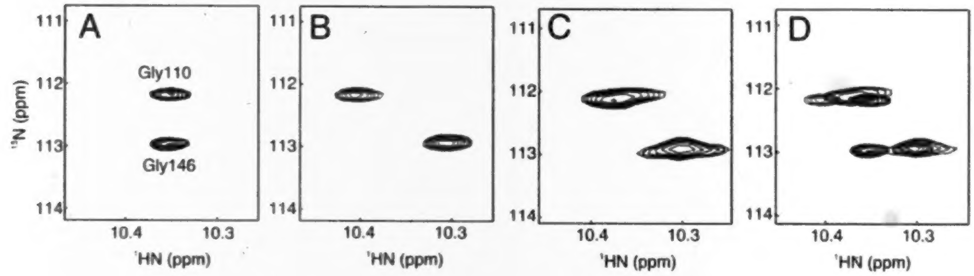


Figure VII-4. **$\{^1H, ^{15}N\}$-HSQC of ^{15}N-cTnC binding to troponin subunits – Gly110 and Gly146**. The progressive amide shifts for Gly110 and Gly146 are expanded from that shown in Figure VII-3, showing (A) C-lobe @ 500MHz (black – unbound), (B) C-lobe @ 500MHz bound to cRP40 (red), (C) cTnC @ 800MHz bound to $cTnI_{34-163}$ and $cTnT_{226-271}$ (TROSY – blue) and (D) an overlay of the three spectra.

VII-C. NMR studies of $cTnI_{34-163}$ within a high molecular weight troponin complex

An $\{^1H, ^{15}N\}$-HSQC NMR spectrum of a complex containing $^{13}C/^{15}N$-$cTnI_{34-163}$ was obtained at 800MHz to determine the extent of spectral dispersion (see Figure VII-5A). The sample revealed a poorly dispersed spectrum with the majority of amide resonances having a ^1H-NMR chemical shift between 7.7ppm to 8.5ppm. As observed in Figure VII-5B, deuteration gave a significant increase in the resolution of individual amide resonances (both samples had identical $^1H/^{15}N$-NMR chemical shifts). A further decrease in line-width was also observed when acquiring in TROSY mode with the acquisition time extended by 4 fold (Figure VII-5C). There were only 87 peaks observed in the 2D-{^{1}H, ^{15}N}-HSQC spectrum, when there should have been 125 backbone resonances for cTnI$_{34\text{-}163}$ (*129 residues – 2 Proline – 1 NH$_3^+$ termini – 1 =* 125). 18 predicted cross peaks corresponding to Asn (1) and Gln (8) were not observed in the TROSY spectra. Therefore, roughly 1/3 of expected backbone amide resonances were not observed in the {^{1}H, ^{15}N}-HSQC spectrum, indicating that there may be some form of exchange broadening and/or motion occurring within the complex in solution.

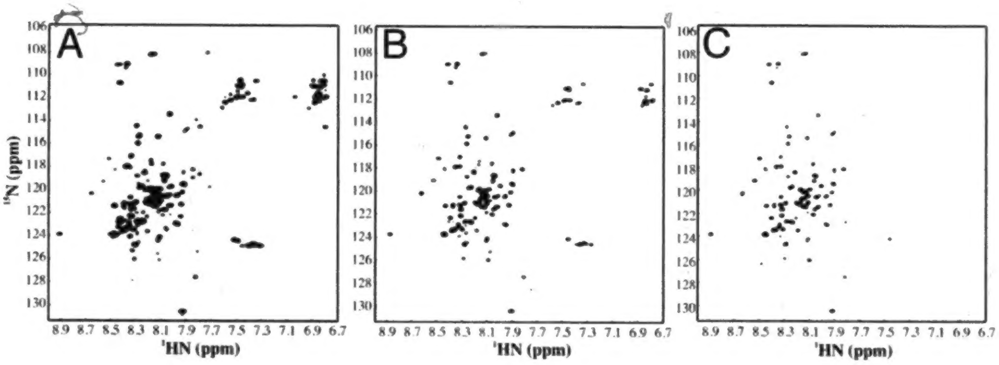


Figure VII-5. **Spectral dispersion of the cTnI$_{34\text{-}163}$ at 800MHz.** ^{15}N-HSQC NMR spectra of (A) non-deuterated ^{13}C/^{15}N-cTnI$_{34\text{-}163}$ and (B) deuterated ^{2}H/^{15}N-cTnI$_{34\text{-}163}$ in a 38.8kDa troponin complex. (C) ^{15}N-HSQCTROSY NMR spectrum of ^{2}H/^{15}N-cTnI$_{34\text{-}163}$.

Comparison of the {^{1}H, ^{15}N}-HSQCTROSY spectrum of cTnI$_{34\text{-}163}$ in the complex with the {^{1}H, ^{15}N}-HSQC spectrum of the cIp region (cTnI$_{128\text{-}147}$) bound to the C-lobe of cTnC reveal differences in the chemical shifts of backbone resonances (Figure VII-6). It is not surprising in that the cIp region is bound only to the C-lobe of cTnC (Chapter III), whereas cTnI$_{34\text{-}163}$ is bound to both cTnC and cTnT$_{226\text{-}271}$. As well, the first eight residues of cIp (cTnI$_{128\text{-}136}$) are expected to participate in a coiled-coil motif with cTnT as per the Takeda crystal structure (see Figure VII-1), which would affect the amide bond chemical shifts.

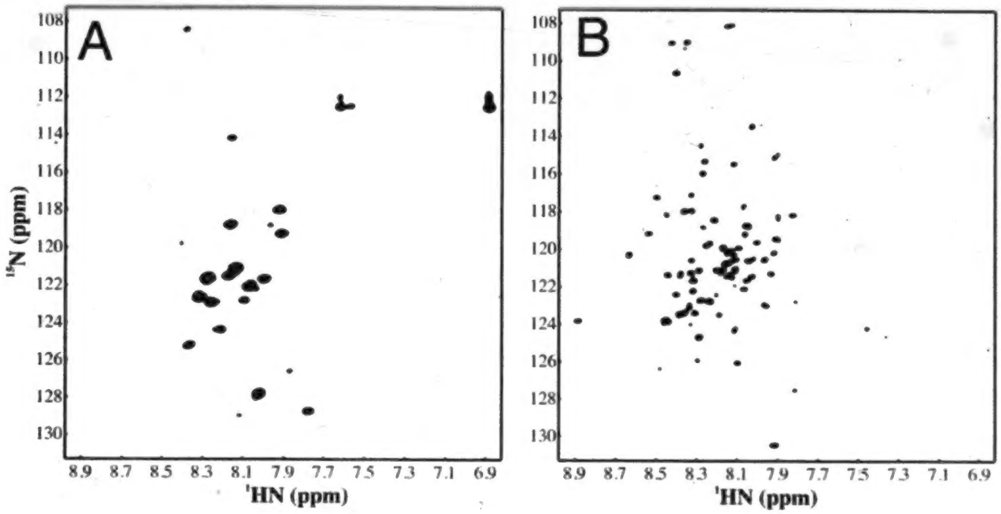


Figure VII-6. **Amide dispersion comparison of the inhibitory region and cTnI$_{34\text{-}163}$.** HSQC spectra of (A) inhibitory region of cTnI$_{128\text{-}147}$ bound to the C-lobe of cTnC at 500MHz and (B) cTnI$_{34\text{-}163}$ bound to the Tn complex (38.8kDa) TROSY at 800MHz.

Elucidation of backbone assignments based solely on the ^{15}N-NMR spectra is difficult as deuteration was required to obtain reasonable line shape, which subsequently removed all carbon-bound proton resonances. Although the majority of cTnI residues are predicted to possess a helical secondary structure (based upon the crystal structure – Figure VII-1), there was a relative lack of d_{NN}'s observed in the noesy-hsqc spectra at 800MHz (data not shown). It was hoped that heteronuclear NMR experiments involving ^{13}C would aid in alleviating ^{15}N overlap issues and resolution, which takes advantage of the chemical shift dispersion for ^{13}C nuclei. ^{13}C-NMR experiments performed on the ^{13}C/^{15}N/^{2}H-cTnI$_{34\text{-}163}$ complex yielded sufficient signal-to-noise for resolution of hnco (500MHz) and hnca (800MHz) experiments. Lower sensitivity ^{13}C-NMR experiments nhcacb and cbca(co)nnh were performed on the ^{13}C/^{15}N-cTnI$_{34\text{-}163}$ sample at 600MHz, however these experiments yielded poor signal to noise, attributed in part to the lack of deuteration for this sample. Due to the relative lack of sensitivity, analysis of the nhcacb and cbca(co)nnh spectra yielded limited usable data.

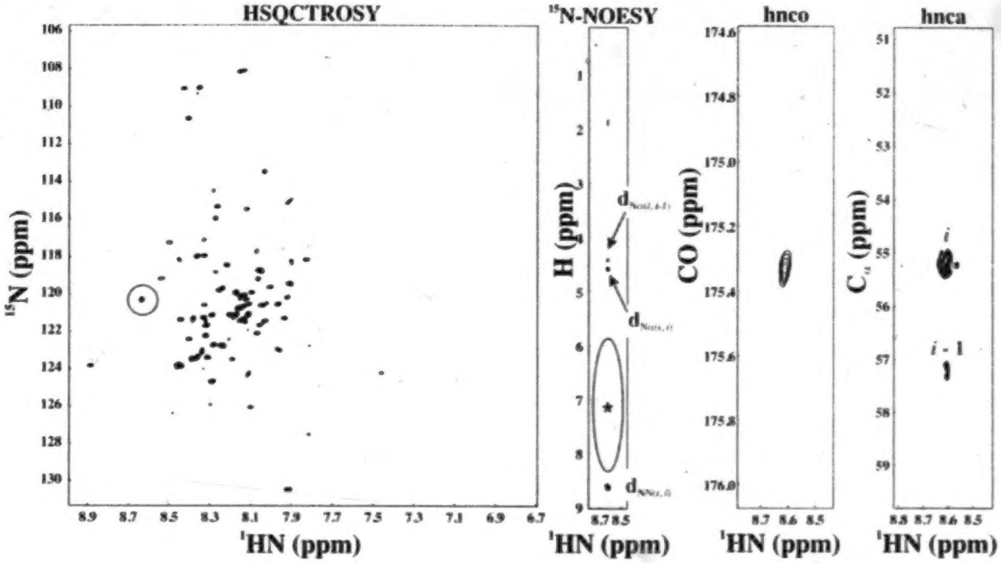


Figure VII-7. **Example of cTnI$_{34\text{-}163}$ amide connectivities**. Connectivity for a 'clean' peak in the hsqcTROSY spectrum, with indicated crosspeaks for the ^{15}N-NOESY, hnco and hnca indicated in red. A red circled area (*) in the ^{15}N-NOESY spectrum indicates a lack of sequential d_{NN} cross peaks.

Figure VII-7 displays an example for connectivities of a 'clean' resonance within the {^{1}H, ^{15}N}-HSQCTROSY spectra, with correlation(s) demonstrated for {^{1}H, ^{15}N}-HSQCTROSY, {^{1}H, ^{15}N}-noesyhsqc, hnco & hnca NMR spectra. The hnco and hnca spectra demonstrated high signal-to-noise, with characteristic strong (*i*, *i*) connectivities observed in both, and a weak (*i*, *i*-1) crosspeak in the hnca. The crosspeak intensity observed in the two ^{13}C/^{15}N experiments was not paralleled in the {^{1}H, ^{15}N}-noesyhsqc, as only correlations of $d_{N\alpha(i,\,i)}$ and $d_{N\alpha(i,\,i-1)}$ were observed. No sequential d_{NN}'s for this resonance are reported (one is observed on the diagonal – $d_{NN(i,\,i)}$), which is mirrored by all resonances in the spectra. The high crosspeak intensities observed in the hnco and hnca spectra were localized only to the outlying resonances in the {^{1}H, ^{15}N}-HSQCTROSY, with those clustering between 118ppm-124ppm demonstrating poor signal intensity. Supplemental sequential assignment ^{13}C/^{15}N-NMR experiments such as the hncacb and cbca(co)nnh also displayed this poor signal-to-noise in this region of the ^{15}N-dimension. It is due to this poor resolution in the ^{15}N-dimension in the acquired ^{15}N- and ^{13}C/^{15}N-NMR experiments that sequential assignment(s) of the protein backbone in the complex are ~95% incomplete, as per the semi-automated assignment program SMARTNOTEBOOK (*43*) and manual inspection.

VII-D. Comparison of skeletal and cardiac TnI within the troponin complex

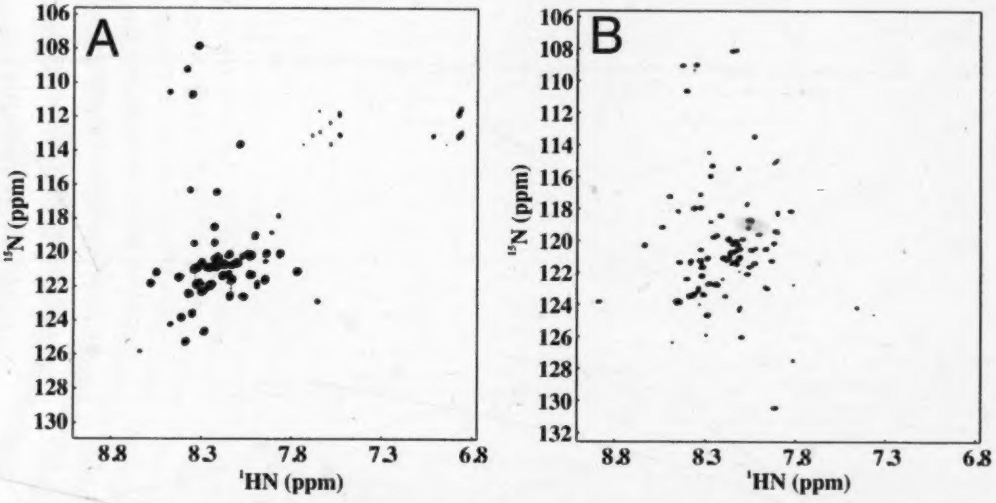


Figure VII-8. **Comparison of skeletal and cardiac TnI in the troponin complex.** {^{1}H, ^{15}N}-HSQCTROSY NMR spectroscopy for (A) skeletal TnI (55kDa) at 600MHz and (B) cardiac TnI (38.8kDa) at 800MHz.

Investigations into TnI have been initiated in the Sykes laboratory by post-doctoral fellow Tharin M. Blumenschein using the skeletal isoform which closely parallel my cardiac studies. She has been investigating a skeletal complex using ^{2}H/^{13}C/^{15}N-sTnI$_{1-182}$ in a 55kDa complex. A direct comparison of the 2D {^{1}H, ^{15}N}-HSQC spectrum of the skeletal isoform with that of cardiac is shown in Figure VII-8. Of all TnI residues in the skeletal complex, 45 possessed strong intensities (see Figure VII-8A), which have been completely assigned and correspond to residues in the region sTnI$_{137-182}$. These assigned resonances correspond to the C-terminal residues of the cardiac isoform, which I truncated during cloning (cTnI$_{170-210}$), and were thus not observed in my cardiac spectra. It is inferred from ^{15}N-NMR relaxation measurements that these skeletal C-terminal residues may form a separate domain not seen in the crystal structures, possessing a greater degree of rotation from the core complex, allowing for the more intense, sharp lines observed in Figure VII-8A. This domain may be occurring in the cardiac isoform as well. Lowering the contour level in the NMR spectrum of the skeletal isoform a number of broader lines below the sharp resonances, which are not currently assigned and it is assumed that they correspond to regions of sTnI that participate in the formation of the core troponin complex. Due to these differences a direct comparison of the ^{15}N-hsqcTROSY spectra is not be completely applicable; however, a certain commonality can be established for the two isoforms. Specifically, both isoforms display a large degree of localized spectral overlap in the ^{15}N-dimension (118ppm-124ppm), which suggests a large degree of helical secondary structure. The differences in the spectra of the two isoforms are the differing molecular weights of the complexes (55kDa vs. 38.8kDa) and the differences in field strength for TROSY acquisition (600MHz vs. 800MHz).

VII-E. ^{15}N-T_2 NMR relaxation studies of cTnI in the troponin complex

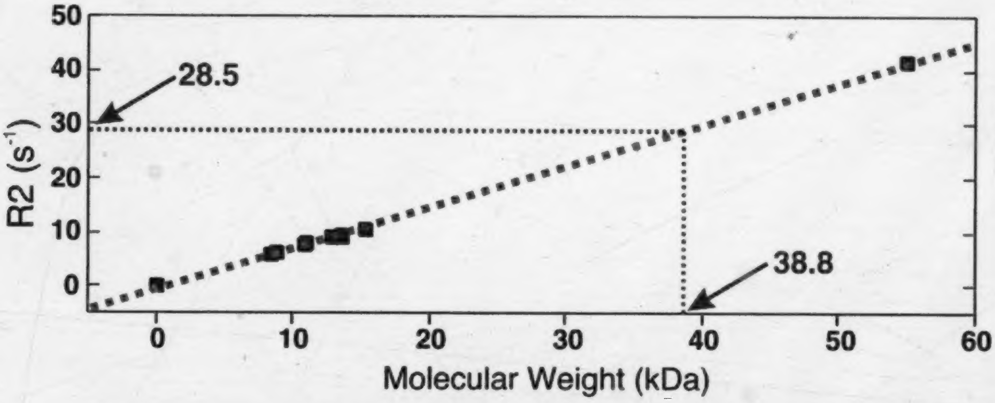


Figure VII-9. **^{15}N-R2 versus molecular weight for cTnI$_{34\text{-}163}$.** An R2 value of 28.5s^{-1} is predicted for cTnI$_{34\text{-}163}$ in the 38.8kDa complex. All data points are taken from relaxation values reported in Chapter V, plus one additional value from Tharin Blumenschein for a 54kDa sTnC complex based upon ^{15}N-T_2 envelope measurements.

The difficulty in resonance assignments for cTnI$_{34\text{-}163}$ in complex was further exemplified during ^{15}N-NMR relaxation measurements. Peak broadening of amide resonances became extremely troublesome for all resonances as linear prediction in the ^{15}N-dimension during spectral processing was omitted, and acquisition of ^{15}N-T_2

relaxation values in TROSY mode was not an available option. Based upon the linear relationship of R2 versus molecular weight previously established in Chapter V of this thesis (Figure V-5), a T_2 average relaxation value for all backbone amide resonances of cTnI in the cardiac complex is expected to be approximately 35.1ms (i.e. R2=28.5s^{-1}), if the complex tumbled as a single rigid isotropic body (Figure VII-9).

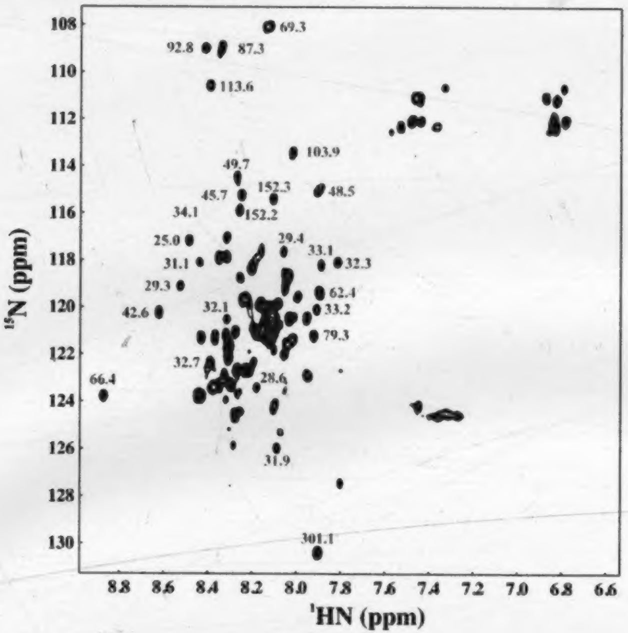


Figure VII-10. **^{15}N-T_2 NMR relaxation values for cTnI$_{34\text{-}163}$ in the troponin complex**. Values are reported in red as milliseconds for outlying residues within the spectra.

The ^{15}N-R2 relaxation values obtained for cTnI$_{34\text{-}163}$ in the complex were difficult to obtain due to peak overlap especially since no linear prediction could be used during processing. ^{15}N-R2 relaxation measurements were obtained for both ^{13}C/^{15}N- and ^{2}H/^{15}N-cTnI$_{34\text{-}163}$ samples; however, only relaxation values for the deuterated sample are presented as it displayed better resolution than the non-deuterated sample. Figure VII-10 displays the first point of the ^{15}N-R2 relaxation series (no delay), revealing an average T_2 value of 91.3 ± 54.9ms when omitting Asn and Gln side chain resonances. There is a large amount of variation among the reported T_2 values, with some resonances displaying values similar to the predicted result, while some differ by order of magnitude. These outlying resonances, that might be assumed to come from the more structured regions, yield T_2 values that are close to the predicted value. Some of the peaks may correspond to resonances that are in more flexible regions of $cTnI_{34-163}$ and will thereby display higher relaxation values than non-flexible regions.

DISCUSSION

Based upon the data presented for $cTnI_{34-163}$ in the 38.8kDa complex, it is evident this project requires additional time and experimental analysis to reach completeness. It has been demonstrated that the region $cTnT_{197-225}$ is not a necessary requirement for the production of a ternary complex, as it displays no affinity and is positioned away from the core complex in the Takeda crystal structure (1J1E.pdb) (see Figure VII-1) (*30*). The C-terminal cTnI extension beyond the switch region ($cTnI_{147-163}$) does not seem to be required for proper folding of the complex in solution. The N-terminal PKA phosphorylation extension for the cardiac isoform ($cTnI_{1-33}$) was not included in the cloning of the troponin I protein, however recent unpublished data by Dr. Monica Li in the Sykes laboratory strongly supports an interactive effect with the N-lobe of cTnC. This is however not surprising, as this domain is known to be protein kinase A (PKA) phosphorylation dependant and has been shown to possess a marked regulatory effect on cardiac contractile activity in response to β-adrenergic stimulation (*44-47*). Perhaps in future investigations, this domain could be included to monitor changes in calcium sensitivity and/or ternary domain orientation(s) in response to a PKA phosphorylation event such as Ser22 and Ser23.

This chapter involving the full-length troponin complex represented a deviation from other chapters in that, unlike past studies of domain subunits, did not produce high quality NMR spectra and possessed solubility issues. Although troubleshooting helped solve the majority of problems, many issues still require further investigation. The sample concentration could only be increased with increasing the buffer ionic strength to non-physiological levels. An increase of sample concentration to 1mM (>3-fold) would allow for greater signal to noise in the lower sensitivity ^{13}C-NMR experiments (i.e. cbca(co)nnh) and would greatly aid in reduced spectrometer time limitations, saving time and money. Investigation of the published crystal structures coordinates may yield non-critical aliphatic residues that could be mutated to ionic to increase the solubility of the complex (*30, 33*). The reported ^{15}N-R2 value for TnC in the skeletal complex by T. Blumenschein closely fit the linear line of best fit offered by the smaller, globular, more isotropic samples it accompanied, yet can a direct comparison between a relaxation value for sTnC and cTnI, individually complexed, be established? There appears to be a yes/no answer, as although considered reasonably similar in implied function, the static crystal structures reveal a differing story. Figure VII-11 reveals the striking differences between the skeletal and cardiac crystal structures, with the cardiac isoform possessing a flexible D/E TnC helix region while the skeletal isoform reports this region as non-flexible. The skeletal structure appears to possess a more rigid, spherical profile than the extended cardiac structure. This observation provides the possibility that the skeletal isoform may behave more isotropic in solution, thus a good ^{15}N-R2 correlation was observed. The large variations in the cTnI R2 values may be due to a differing rotational axis, resulting in potential line broadening due to exchange of flexible regions.

The majority of proposed differences between the two structures occur within the region starting C-terminal from the inhibitory region, and the head orientations of TnC (see Figure VII-11). The coiled-coil motif observed between the TnT and TnI subunits for both structures are quite similar which reveals the structural importance of this IT domain. Additionally, the C-lobe of TnC associates with this coiled-coil domain in a similar manner for both structures. Coiled-coil motifs are traditionally quite stable and have been used in industry as biosensors and affinity purification tools. Structural studies of coiled-coil motifs by NMR has traditionally been extremely difficult and troublesome, as the repeating heptad possesses relatively little variability in amino acid sequence and chemical shift, making for large overlap in the NMR assignment process (for an example, see Appendices F & G, where the solution structure of a heterodimeric EK coiled-coil was elucidated as part of a collaboration with Robert S. Hodges).

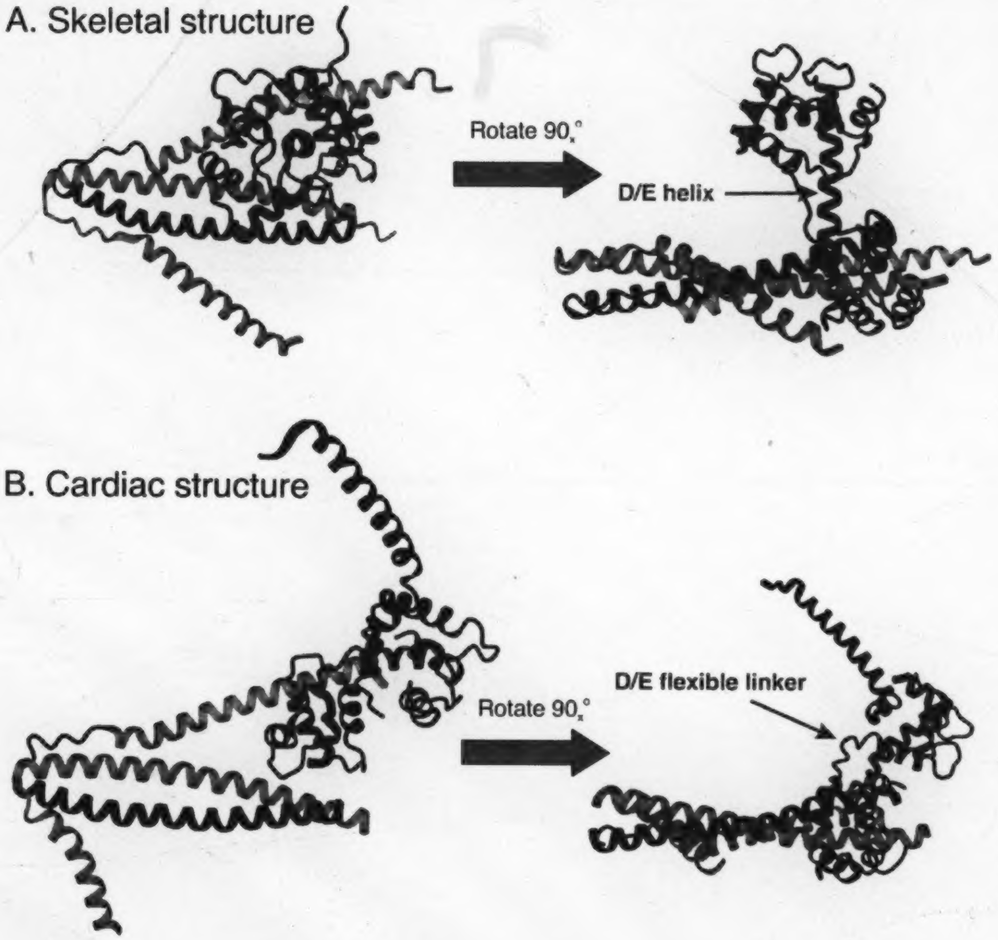
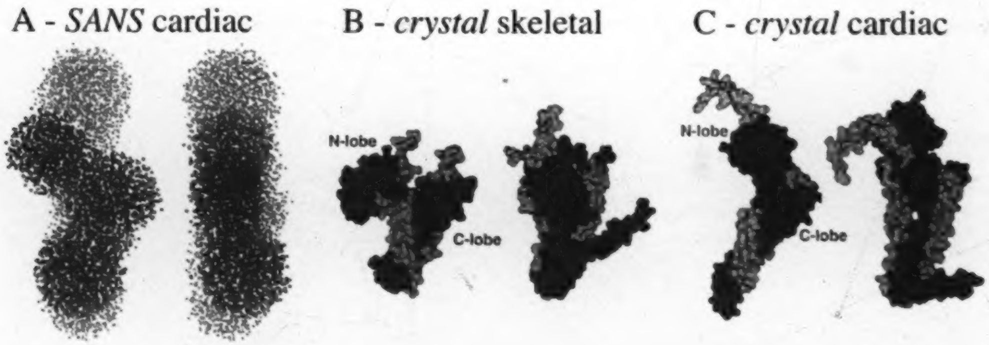


Figure VII-11. **Comparison of cardiac and skeletal crystal structures**. Ribbon diagrams of the crystal structures of the (A) skeletal (1YTZ.pdb) and (B) cardiac (1J1E.pdb) isoforms shown looking down the TnC E-helix (left) and a 90° rotation (right) (*30, 33*). The cardiac isoform reveals a more extended structure than that for the skeletal, due to a flexible D/E linker region within TnC. Structures are colored as: TnC (red), TnI (cyan/blue) and TnT (green).

Figure VII-12 reveals a striking difference in domain orientations investigation of the cardiac SANS structure (cTnC•cTnI•cTnT$_{198\text{-}298}$) reported by Heller when compared to the Takeda cardiac crystal structure (1J1E.pdb) (*17, 30*). The cardiac SANS structure reported TnC head orientations that closely resemble the skeletal crystal structure reported by Vinogradova (*33*). The cardiac crystal structure has an extended, open conformation that may simply be an artifact of crystal packing. This comparison strongly suggests that the preferred solution conformation of the intact troponin complex adopts an orientation similar to the cardiac SANS structure and not the cardiac crystal structure. This conclusion is far reaching with physiological potentials, as it appears that there is finally a consensus on the preferred structure of the troponin complex. Importantly, the crystal structure of the skeletal isoform appears to behave as would be expected for it in solution (SANS structure). More importantly, my reported studies involving cTnI, cTnT and cTnT through out this thesis strongly supports the cardiac SANS and skeletal crystal structure.



A - *SANS* cardiac B - *crystal* skeletal C - *crystal* cardiac

Figure VII-12. **Domain orientations of the troponin complex**. Domain orientation of the solved troponin isoform complexes by (A) SANS cardiac, (B) crystal skeletal (1YTZ.pdb) and (C) crystal cardiac (1J1E.pdb) (*17, 30, 33*). Panels B & C are oriented in such a way that they match the SANS structure in panel A, with superimposition of the C-lobe of TnC. For all panels, TnC/TnT are red and TnI is blue. The cardiac SANS structure more closely resembles the skeletal crystal structure. Panel A is taken from (*17*).

Studies of the intact cardiac troponin complex have revealed a strong similarity to that the skeletal isoform, which is not altogether shocking, as these two isoforms perform a similar role in muscular contraction and relaxation events within the body. Yet the simple fact that there is two different isoforms for two differing types of tissues (skeletal vs. cardiac) dictates that differences, however minute, are physiologically important. Can structural inferences from the skeletal isoform explain physiological results observed for the cardiac isoform? The intention of the work performed in this chapter was an attempt to answer, in part, this question. Unfortunately since this work is not yet complete, a definitive answer is not yet possible. A concluding discussion regarding the structural consequences of the cardiac and skeletal isoforms on physiological functionality, such as response to phosphorylation on activity, will be further expanded in Chapter VIII. Does the solution structure dictate the physiological happenings in the thin filament, or is another orientation possible? All of these questions and more concluding discussions will be addressed in the concluding chapter of this thesis: Chapter VIII.

REFERENCES

1. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.
2. Lindhout, D. A., Litowski, J. R., Mercier, P., Hodges, R. S., and Sykes, B. D. (2004) *Biopolymers 75*, 367-75.
3. Lindhout, D. A., and Sykes, B. D. (2003) *J Biol Chem 278*, 27024-34.
4. Wang, X., Li, M. X., and Sykes, B. D. (2002) *J Biol Chem 277*, 31124-33.
5. Wang, X., Li, M. X., Spyracopoulos, L., Beier, N., Chandra, M., Solaro, R. J., and Sykes, B. D. (2001) *J Biol Chem 276*, 25456-66.
6. Li, M. X., Wang, X., Lindhout, D. A., Buscemi, N., Van Eyk, J. E., and Sykes, B. D. (2003) *Biochemistry 42*, 14460-8.
7. Li, Y., Love, M. L., Putkey, J. A., and Cohen, C. (2000) *Proc. Natl. Acad. Sci. U. S. A. 97*, 5140-5.
8. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.
9. Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) *FEBS Lett 453*, 107-12.
10. Finley, N., Dvoretsky, A., and Rosevear, P. R. (2000) *J Mol Cell Cardiol 32*, 1439-46.
11. Finley, N. L., Howarth, J. W., and Rosevear, P. R. (2004) *Biochemistry 43*, 11371-9.
12. Mercier, P., Spyracopoulos, L., and Sykes, B. D. (2001) *Biochemistry 40*, 10063-77.
13. Mercier, P., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 2902-11.
14. Brown, L. J., Sale, K. L., Hills, R., Rouviere, C., Song, L., Zhang, X., and Fajer, P. G. (2002) *Proc Natl Acad Sci U S A 99*, 12765-70.
15. Dvoretsky, A., Abusamhadneh, E. M., Howarth, J. W., and Rosevear, P. R. (2002) *J Biol Chem 277*, 38565-70.
16. Heller, W. T., Abusamhadneh, E., Finley, N., Rosevear, P. R., and Trewhella, J. (2002) *Biochemistry 41*, 15654-63.
17. Heller, W. T., Finley, N. L., Dong, W. J., Timmins, P., Cheung, H. C., Rosevear, P. R., and Trewhella, J. (2003) *Biochemistry 42*, 7790-7800.
18. Hernandez, G., Blumenthal, D. K., Kennedy, M. A., Unkefer, C. J., and Trewhella, J. (1999) *Biochemistry 38*, 6911-7.
19. Tung, C. S., Wall, M. E., Gallagher, S. C., and Trewhella, J. (2000) *Protein Sci 9*, 1312-26.
20. Tripet, B., De Crescenzo, G., Grothe, S., O'Connor-McCourt, M., and Hodges, R. (2002) *J Mol Biol 323*, 345.
21. Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) *J Mol Biol 271*, 728-50.
22. McKay, R. T., Saltibus, L. F., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 12731-12738.
23. McKay, R. T., Pearlstone, J. R., Corson, D. C., Gagne, S. M., Smillie, L. B., and Sykes, B. D. (1998) *Biochemistry 37*, 12419-30.

24. Gagne, S. M., Li, M. X., McKay, R. T., and Sykes, B. D. (1998) *Biochem Cell Biol 76*, 302-12.
25. Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry 36*, 4386-4392.
26. Gagné, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) *Prot. Sci. 3*, 1961-1974.
27. Gagné, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) *Nature Struct. Biol. 2*, 784-789.
28. Putkey, J. A., Sweeney, H. L., and Campbell, S. T. (1989) *J. Biol. Chem. 264*, 12370-8.
29. Putkey, J. A., Liu, W., and Sweeny, H. L. (1991) *J. Biol. Chem. 266*, 14881-14884.
30. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
31. Matsumoto, F., Makino, K., Maeda, K., Patzelt, H., Maeda, Y., and Fujiwara, S. (2004) *J Mol Biol 342*, 1209-21.
32. King, W. A., Stone, D. B., Timmins, P. A., Narayanan, T., von Brasch, A. A., Mendelson, R. A., and Curmi, P. M. (2005) *J Mol Biol 345*, 797-815.
33. Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) *Proc Natl Acad Sci U S A*.
34. Riek, R., Wider, G., Pervushin, K., and Wüthrich, K. (1999) *Proceedings of the National Academy of Sciences 96*, 4918-4923.
35. Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997) *Proceedings of the National Academy of Sciences 94*, 12366-12371.
36. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
37. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603-614.
38. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.
39. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.
40. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984-6003.
41. Ikura, M., Kay, L. E., and Bax, A. (1990) *Biochemistry 29*, 4659-67.
42. Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson. B103*, 203-216.
43. Slupsky, C. M., Boyko, R. F., Booth, V. K., and Sykes, B. D. (2003) *J Biomol NMR 27*, 313-21.
44. Smith, J. J. (1980) *Circulatory Physiology: The Essentials*, Baltimore, Md: Waverly Press.
45. Solaro, R. J., Moir, A. J., and Perry, S. V. (1976) *Nature 262*, 615-7.
46. Katz, A. M. (1990) *Eur Heart J 11 Suppl A*, 27-31.
47. Metzger, J. M., and Westfall, M. V. (2004) *Circ Res 94*, 146-58.

CHAPTER VIII

Conclusion

OVERVIEW

This chapter provides a concluding discussion regarding the state of cardiac troponin I. It will begin by providing a brief summary of the nature of the work reported in Chapters II to VII of this thesis, and build upon the conclusions reached in these chapters to include structural and physiological implications for the troponin complex in cardiocytes. Prior conclusions regarding mutational and phosphorylation studies of troponin I will be expanded to include my current view(s) of the state of the troponin complex; where is has progressed from and where we are currently. Particular emphasis will be placed on the importance for understanding diseased states at the atomic level and the regulation (via phosphorylation) of thin filament contraction, proving that troponin I is without doubt the 'I' in team.

SUMMARY

This thesis was initiated with the intent of enhancing our understanding of the structural implications for cardiac troponin I (cTnI) within the troponin complex. Intricately woven with binding partners troponin C (cTnC) and T (cTnT), these proteins are the primary ignition source for cardiac contraction via a calcium-signaling cascade within the sarcomere. Troponin I is the sole protein responsible for cardiac relaxation/contraction due to its ability to toggle on/off the myosin binding site on actin within the thin filament. Possessing multiple domains, cTnI has one critical 20-residue section that is responsible for the inhibitive activity: this basic stretch is referred to as the inhibitory region (cIp). Obtaining the structure, location and dynamics of cIp during the toggling on/off of actin during contraction/relaxation cycles is of utmost importance in the goal of understanding muscular contraction at the atomic level. For this reason I chose to elucidate the NMR solution structure of the cIp region bound to its contractory binding partner cTnC.

Elucidation of the binary structure of the inhibitory region bound to the C-lobe of cTnC required the development of a system to produce high quantities of labeled (i.e. ^{13}C and/or ^{15}N) recombinant peptide(s). Using a 56 amino acid fusion protein expression construct, protein G from the *Streptococcus* B1 domain, I was able to produce the necessary quantities of $^{13}C/^{15}N$-cIp required for the production of the NMR solution structure of a binary inhibited complex (Chapter II). This inhibited complex revealed that cIp binds rigidly across the hydrophobic face of the C-lobe, making numerous electrostatic interactions with the E-helix of the C-lobe (Chapter III). Electrostatic analysis of these interactions was mapped using scanning glycine mutagenesis, proving the nature of association between the inhibitory region and the C-lobe was dictated via charge distribution and placement (Chapter IV). Interestingly, diseased cardiac states such as familial hypertrophic cardiomyopathy (FHC) can arise when mutation(s) of basic residues to non-charged residues occur within cTnI, altering the delicate balance of electrostatic attractions and repulsions within the troponin complex. Additionally, the ability to introduce negatively charged phosphate groups, via phosphorylation by various cellular kinases, is a potent form of fine-control in the regulatory processes of cardiac tissues in response to events such as stress. Consequently, the FHC mutation R144G and the phosphorylation event on Thr142 within the inhibitory region had marked effects on the binding affinity with the C-lobe; lowering the dissociation constant by greater than one order of magnitude. These results conclude that the inhibitory region is a critical region in the basal functioning of contraction, as is evidenced by the high degree of sequence conservation throughout isoforms/species and inability to tolerate mutation(s).

Although critically important to function, the inhibitory region is not the only region of troponin I that makes interactions with other troponin proteins. Troponin C interactions with both the inhibitory region and the N-terminal cRP40 region of cTnI revealed that both regions are capable of binding simultaneously to the C-lobe (Chapter V): results that support the skeletal crystal structure, while drawing into question aspects of the cardiac crystal structure. Interactions of the inhibitory region with the N-lobe of cTnC were observed (Chapter VI), indicating that Arg147 is critical for proper placement and contacts of cIp with the N-lobe. Phosphorylation studies of the switch and cRP40 regions of cTnI (Ser41, Ser43, Ser149) agree with the conclusion reached for Thr142 in that the increased negative charge density of the phosphate group plays a dominant role in the electrostatic environment of the complex, giving these PKC and PAK sites an efficient regulatory control mechanism for the fine-tuning of cardiac contraction. NMR relaxation studies of a larger ternary troponin complex suggest that structure and behavior in solution more closely resembles that of the skeletal crystal structure than that of the cardiac structure (Chapter VII).

DISCUSSION

It is the year 2005: do we really know anything more than what was proposed by Huxley 50 years ago (*1-3*)? I would definitely say yes. We now know most of the protein components that compose the thick and thin filaments. We possess the ability to quickly sequence genes, which greatly aids in the understanding of where and how a single mutation may lead to a mal-functioning protein resulting in a diseased state. We have characterized some of the physical attributes of these diseased states (i.e. patient presentation of symptoms) and can recommend therapies to alleviate symptoms. We have high-resolution structural data of many of these proteins, and can attempt to understand the consequences of mutational effects at the atomic level.

The primary goal of this thesis was to offer structural insight(s) into the workings of the troponin complex during cardiac contraction. Quite simply; do we understand what is really going on within the sarcomere during a contraction event? I feel that this question has a yes/no answer. Based upon direct observation such as electron microscopy and force transduction techniques (*4*, *5*), we know that when muscle tissue is active the individual thin and thick filaments slide past one another in an organized fashion, exactly as was predicted over 50 years ago by Huxley. Thus I can say with relative certainty that we know the how the macromolecular environment responds to a stimuli event. However, the inner workings of the micromolecular environment are not quite as well understood. Structural curiosities of altered protein-protein interactions at the molecular level have yielded a large wealth of atomic resolution data, presenting insights into the force producing machinery (*6-34*). Yet I feel that additional research is still required for a definitive answer in the micromolecular architecture not only for the troponin complex but for all protein interactions within muscle fiber.

So what do we actually know about the structural details of cardiac contraction and regulation, with regards to troponin I? Detailed in Chapter I, the calcium signaling event for muscle contraction is sensed by the low-affinity N-lobe (*13*). Upon binding a single Ca^{2+} ion, the structural consequences are propagated from cTnC to the troponin complex, to the thin filament, and ultimately through the entire sarcomere. The switch region re-locates to the hydrophobic binding face on the N-lobe upon association with calcium, triggering a movement of the inhibitory region off the actin-binding site to interactions with the C-lobe (D/E linker). The data presented in this thesis (Chapter III) proved that the inhibitory region binds rigidly to the C-lobe along the entirely of its backbone, yet the cardiac crystal structure implied that the inhibitory region is away from the C-lobe in the calcium saturated state. NMR analysis of this interaction was addressed using scanning glycine mutagenesis, resulting in a mutual electrostatic interaction between the C-lobe and the inhibitory region. The skeletal crystal structure was in agreement with my reported NMR result: *the inhibitory region interacts with the D/E helix of TnC*.

Functional implications regarding the structural placement of the inhibitory region becomes immediately clear when mutational data is considered. The cardiac disease Familial Hypertrophic Cardiomyopathy (FHC) occurs within individuals with a point mutation R144G within the inhibitory region. Although I was unable to precisely place this residue in spatial relation to the C-lobe, I was able to report that the affinity was lessened by 6-fold upon incorporation of this FHC mutation. As previously discussed, the skeletal crystal structure closely matches my reported NMR results, suggesting the skeletal crystal structure more resembles the physiological solution structure than that of the cardiac crystal structure (*35, 36*). Investigation regarding the placement of Arg144 (Arg112 according to skeletal numbering) reveals that this residue makes a critically important salt-bridging interaction with Glu94 of TnC (Figure VIII-1). I predict that the same salt-bridge pair is present within the cardiac system. However, one might ask why this mutation is only deleterious in the cardiac isoform. If the skeletal isoform possesses this critical salt-bridge, then why isn't there a corresponding disease of the skeletal tissues, which mirrors the cardiac FHC disease? Perhaps the skeletal genes simply do not contain this mutation and therefore it is not inherited as is the case for the cardiac genes, thus it simply does not show up in the human genome with the higher prevalence observed for the cardiac. Another reason may be that skeletal isoform, possessing slightly differing physical architecture than that observed for the cardiac (i.e. longer sarcomeric length, different physiological function), is better able to accommodate a mutation in this region, and thus it just simply does not manifest as a diseased state. Maybe this mutation simply hasn't been genotyped, and has thus gone unnoticed at present. Perhaps the mutation affects interactions with the actin-binding site, thereby altering interactions with actin/tropomyosin during relaxation.

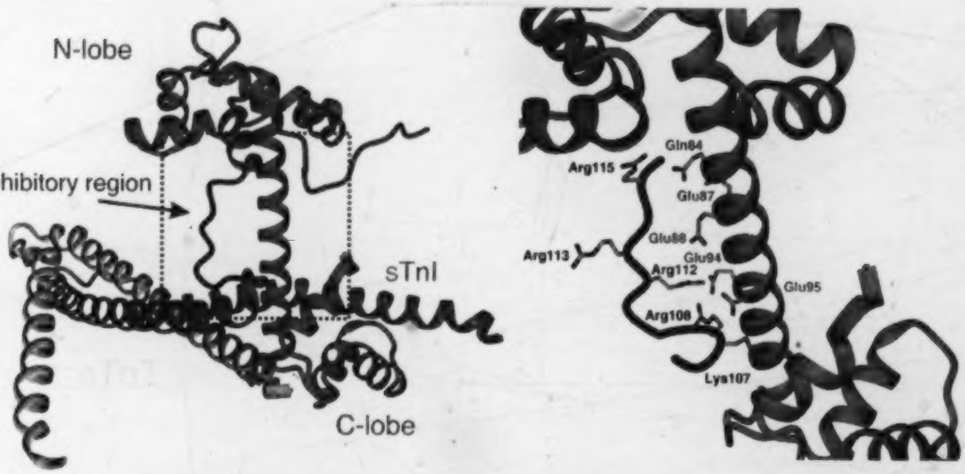


Figure VIII-1. **Identification of FHC mutation R144G on the inhibitory region**. Ribbon diagram of the skeletal crystal structure, where Arg112 (corresponding to Arg144 in the cardiac isoform) participates in a salt-bridge with Glu94 of TnC.

Although the former four possibilities are plausible explanations, I wonder if it might be much more subtle than anyone has previously proposed. We know that the cardiac isoform contains an N-terminal PKA phosphorylation domain that the skeletal isoform does not contain, as well a PKC phosphorylation site at Thr142 that is cardiac specific. We know that phosphorylation of these sites are important in the day-to-day regulatory activity of cardiac tissue. Based upon the crystal structure, it has been proposed that Ser22/Ser23 may be in close vicinity to the inhibitory region on the D/E helix of TnC. Thr142 is only a few angstroms distance from Arg144. Perhaps Arg144 plays an additional role of stabilizing the troponin complex in the Ser22/Ser23 (PKA). More importantly, the FHC mutation abolishes the PKC consensus sequence (Ser/Thr-X-Lys/Arg), therefore the functional implications associated with the diseased FHC state R144G may more be a consequence of incorrect regulation via phosphorylation (Thr142) than altered protein-protein contacts as per the residue Arg144. These phosphorylation pathways are important regulators of cardiac activity throughout the lifetime of an individual, the perpetually altered activity of cardiac regulation is responsible for the left ventricular thickening throughout the lifetime of an individual, eventually leading to premature death. Regardless of the proposed possibilities, I propose that it is a resultant electrostatic imbalance in the cardiac thin filament that is responsible for this altered activity, and not a shift in secondary, ternary and/or quaternary structure.

Given the functional importance of troponin I, it is not surprising that nature has chosen this protein as an effector of cardiac activity via multiple phosphorylation sites. I have demonstrated and discussed in this thesis the effects of three of these modulator kinases: PKC (Thr142, Ser41, Ser43), PAK (Ser149) and PKA (Ser22, Ser23). Troponin I makes large movements on and off both actin and TnC during the cardiac relaxation/contraction cycles, so it makes sense that this troponin subunit is the target of regulatory kinases. Perhaps it has been targeted by kinases due to its movement and flexibility during contraction/relaxation. A domain that is easily accessible to the globular kinases makes a more attractive target than one that might be buried within the thin filament.

The majority of this discussion is focused on the skeletal crystal structure, yet how do we really know that this is really the correct structure *in vivo*? Throughout my thesis, I have been eluding to the fact that troponin I (and the rest of the complex) must be a dynamic entity, therefore are my conclusions based upon a single crystal structure valid? The cardiac crystal structure does not closely resemble my NMR results for the inhibitory region as well it does not correlate well with the skeletal crystal structure and the cardiac small angle neutron scattering data (SANS) (*21, 22, 35, 36*). The cardiac crystal structure appears to have an extended conformation for TnC, whereas the skeletal crystal reports it as more closed. How can I verify which is correct? If the cardiac isoform was locked into a low-temperature conformation, then how I can assume that the skeletal structure doesn't possess the same limitations? There is no simple answer to this question. Obviously the troponin complex must possess the ability to move during contraction, as if it were locked into a single conformation no contraction would be possible. Troponin orientation *in situ* has been investigated using rhodamine fluorescent for the skeletal isoform, establishing functional angles in various physiological states (*37*). These data heavily support the hypothesis of movement and domain re-orientation upon calcium induced contraction events. Personally, I think that the use of the crystal structures is a good idea as a structural scaffold, but this is as far as it should be taken. Understanding muscle physiology cannot be complete based solely upon these structures: we must incorporate dynamic measurements to begin to comprehend the movements during contraction, hence the need for the completion of Chapter VII.

Can the book be closed on the role of troponin I in contraction? Do we understand the structural consequences of this protein in the muscle machinery with respect to other species and tissues? I thought it might be prudent to look into a non-human isoform(s) of troponin I to properly discuss this issue. The invertebrate waterbug *Lethocerus* is well characterized with regards to protein sequence and muscle physiology, and is an appropriate candidate to compare against the cardiac system. Similar to vertebrate skeletal muscle, the regulatory complex of thin filaments in insect flight muscle consists of tropomyosin and three troponin components. Tropomyosin and troponin C are similar in structure and function to vertebrate. Troponin T has a higher molecular weight because of an insertion. Troponin I is not detectable in *Lethocerus* flight muscle (*38*). It is replaced by the 484-residue component troponin H (LTnH), which does not show the same characteristics as vertebrate TnI, instead possessing a large poly-Pro/Ala insertion at the C-terminal region of TnI. This 'abnormal' TnI protein displays a sequence alignment that reveals a poor conservation between the itself and the cardiac troponin I isoform, with the exception of the inhibitory region. The conservation of the inhibitory region is not surprising as this region possesses the critical residues for inhibitive activity. These critical basic residues responsible for electrostatic interactions are highly conserved, again pointing to the intolerance of mutation(s) in this region (*39*). Strangely, the switch region in LTnH does not resemble the cardiac isoform, suggesting that it does not interact with the N-lobe in a calcium-dependant manner during contraction.

Sequence alignment – LTnH and cTnI

Protein 1: *Lethocerus* flight protein Troponin H – 484 residues
Protein 2: Human cardiac troponin I – 210 residues

```
Protein 1: MADEEAKKAKQ     AEIERKRAEVRKRMEEASKAKKAKKGFMTPERKK   45
Matching.: |||  *  |**     | | |* *  |    |    ||*|       ||
Protein 2: MADGSSDAAREPRPAPAPIRRRSSNYRAYATEPHAKKKSK   ISASRKL   47

Protein 1: KLR LLRKKAAEELKKEQERKAAERRRIIEERCGHCCDVDNANEDDLHEV   94
Matching.:  |* ||   | *|| *| | *  |* |    ||    **      *| **
Protein 2: QLKTLLLQIAKQELEREAEERRGEKGRALSTRCQPL ELAGLGFAELQDL   96

Protein 1: IQHYYDRMYTCESQKWDLEYEVRKRDLEISDLNSQVNDLRGKFVKPTLKK   144
Matching.:       |    * ***| |  | |   ||*||   * |||||| *|||**
Protein 2: CRQLHARVDKVDEERYDIEAKVTKNITEIADLTQKIFDLRGKFKRPTLRR   146
                                              Inhibitory region

Protein 1: VSKYENKF  AKLQKKAAE FNFRNQLKVVKKKEFTMEEEEKEKK VVDW   190
Matching.: |    *    | |  *| | ***|  || |||     |* |||** | ||
Protein 2: VRISADAMMQALLGARAKESLDLRAHLKQVKK     EDTEKENREVGDW   191
                Switch region

Protein 1: SKKDEKKVDEEAVVNSDVAELSSVQESVSQEAPPVDTQSEASEIPTTTTQ   240
Matching.:  |*    *|  *        * |
Protein 2: RKN    IDALSGMEGRKKKFES                              210

Protein 1: EEHKQEEQQQQQPPPPPKTEDDLWLFLKEAITSANPMTLEELQQKLIDT   290

Protein 1: VSHVPPDFLEPVVEASIKRSQSMDQLASISSEVAEGGPPAEPGKEGEAAP   340

Protein 1: PAEGAPAAAPPAEGAPPAEGTPAAAAPPAEGAPAAAAPPAEGAPPAEGAA   390

Protein 1: PAAAPPAEGAPAAAAPPAEGAPAAAAPPAEGAQAAAAPPAEGAPAAGAPP   440

Protein 1: AEGAPPAAAAAPPAEGAAPPAEGAAAPAPAIQPVDAPPQVQAAA         484
```

Given that *Lethocerus* flight muscle is capable of a much higher frequency of contraction/relaxation when compared to human cardiac/skeletal isoforms; how important are other troponin I interactions with the rest of the complex? Figure VIII-2 diagrams the interactions of the inhibitory/switch regions of TnI/TnH with TnC. It is very interesting to note that *Lethocerus* only contains a single calcium-binding site (site IV) in the C-terminal lobe, with no calcium affinity observed for the N-lobe (*38*). This thesis has re-enforced the theory that the switch region is critical in initiating the re-location of troponin I off actin, via interactions with the N-lobe in a calcium-dependant manner; however, insect flight muscle doesn't follow this rule. Perhaps the binding of the positive Ca^{2+}-ions within this negatively charged N-lobe masks, in part, the electronegative nature of this domain allowing for a hydrophobic binding event. This logical assumption is based directly on the apparent K_D affinities of the switch isoforms for the N-lobe.

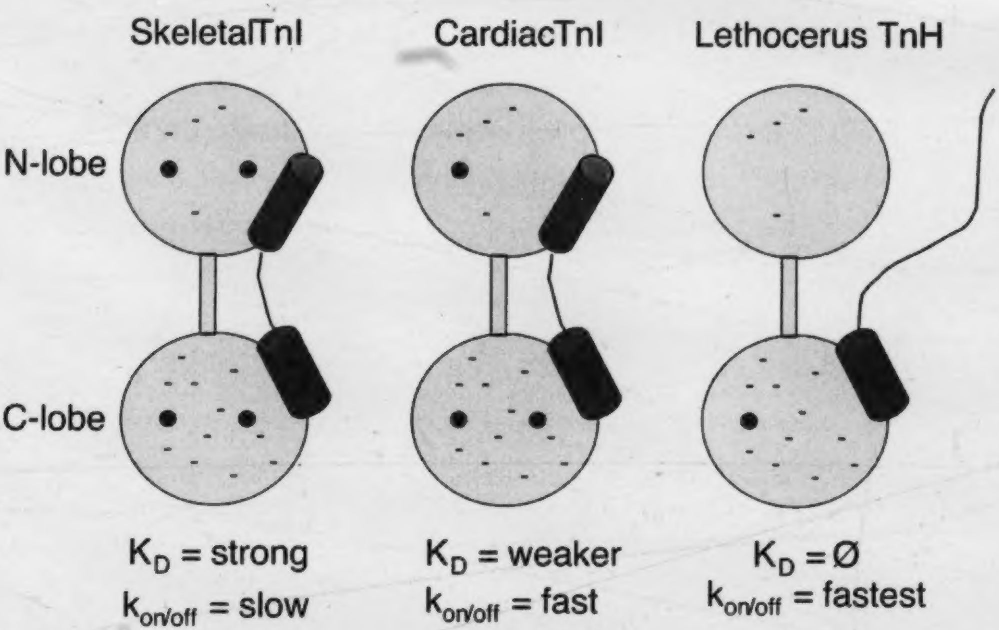


Figure VIII-2. **Comparison of *Lethocerus* troponin H and skeletal/cardiac troponin I with troponin C.** The switch region of TnI is drawn as a blue cylinder, the inhibitory region as a red cylinder and TnC as a yellow dumbbell. Calcium ions are depicted as blue dots, and domain negative charges by (-). The K_D rate displays the affinity of the switch region for the N-lobe for each respective isoform, which are inversely related to the $K_{on/off}$ rate.

The apparent difference(s) in the on/off rates of the inhibitory/switch regions for these isoforms yields the potential for physiological and functional understandings. As LTnH retains a functional inhibitory region, the inhibitive activity of TnH is conserved. Yet since *Lethocerus* flight can function without a switch region, why does it appear to be critical in the human cardiac and skeletal isoforms? Evidently the N-lobe might be important in the regulation of the frequency of contraction, as the affinity of the switch region for the N-lobe may act as a modulator of contraction frequency. Perhaps if humans did not possess the switch region of TnI in our muscular tissues, we would be doomed to potentially catastrophic ultra-high frequency contractions. It is of interest to note that insects such as *Lethocerus* do possess a 'normal' troponin I gene in other muscular tissues (i.e. legs), which do not realize these high frequency contractions. Also, as observed with the cardiac isoform, LTnH possesses the equivalent cardiac Thr142 as a potential phosphorylation site, however the phosphorylation effects on LTnH activity are currently not well characterized. All other studied phosphorylation sites are not conserved in LTnH, with the exception of cTnI-Ser41. The C-terminal poly-Pro/Ala extension is not well understood in regards to functionality. Perhaps it plays an important role in the ability to sustain high-frequency contractions as its bulk may aid as a steric block allowing for the greater frequency observed.

The last 10 years has witnessed a large surge in the structural understanding of the troponin complex, and in particular troponin I. Although subtle, the differences in the sequence identities between the skeletal and cardiac isoforms render them uniquely suited for regulation of their intrinsic tissues. This thesis has investigated and highlighted the critical interactions of the cardiac isoform within the troponin complex by NMR. Although further years of study are required to fully elucidate the entirety of muscular contraction and regulation, it is my hope that the work reported within this thesis has contributed to the greater understanding of muscular contraction.

REFERENCES

1. Hanson, J., and Huxley, H. E. (1953) *Nature 172*, 530-2.
2. Huxley, H. E. (1953) *Biochim Biophys Acta 12*, 387-94.
3. Huxley, H. E. (1953) *Proc R Soc Lond B Biol Sci 141*, 59-62.
4. Perry, S. V. (1999) *Mol. Cell. Biochem. 190*, 9-32.
5. Xu, C., Craig, R., Tobacman, L., Horowitz, R., and Lehman, W. (1999) *Biophys. J. 77*, 985-992.
6. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.
7. Lindhout, D. A., Litowski, J. R., Mercier, P., Hodges, R. S., and Sykes, B. D. (2004) *Biopolymers 75*, 367-75.
8. Lindhout, D. A., and Sykes, B. D. (2003) *J Biol Chem 278*, 27024-34.
9. Wang, X., Li, M. X., and Sykes, B. D. (2002) *J Biol Chem 277*, 31124-33.
10. Wang, X., Li, M. X., Spyracopoulos, L., Beier, N., Chandra, M., Solaro, R. J., and Sykes, B. D. (2001) *J Biol Chem 276*, 25456-66.
11. Li, M. X., Wang, X., Lindhout, D. A., Buscemi, N., Van Eyk, J. E., and Sykes, B. D. (2003) *Biochemistry 42*, 14460-8.
12. Li, Y., Love, M. L., Putkey, J. A., and Cohen, C. (2000) *Proc. Natl. Acad. Sci. U. S. A. 97*, 5140-5.
13. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.
14. Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) *FEBS Lett 453*, 107-12.
15. Finley, N., Dvoretsky, A., and Rosevear, P. R. (2000) *J Mol Cell Cardiol 32*, 1439-46.
16. Finley, N. L., Howarth, J. W., and Rosevear, P. R. (2004) *Biochemistry 43*, 11371-9.
17. Mercier, P., Spyracopoulos, L., and Sykes, B. D. (2001) *Biochemistry 40*, 10063-77.
18. Mercier, P., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 2902-11.
19. Brown, L. J., Sale, K. L., Hills, R., Rouviere, C., Song, L., Zhang, X., and Fajer, P. G. (2002) *Proc Natl Acad Sci U S A 99*, 12765-70.
20. Dvoretsky, A., Abusamhadneh, E. M., Howarth, J. W., and Rosevear, P. R. (2002) *J Biol Chem 277*, 38565-70.
21. Heller, W. T., Abusamhadneh, E., Finley, N., Rosevear, P. R., and Trewhella, J. (2002) *Biochemistry 41*, 15654-63.
22. Heller, W. T., Finley, N. L., Dong, W. J., Timmins, P., Cheung, H. C., Rosevear, P. R., and Trewhella, J. (2003) *Biochemistry 42*, 7790-7800.
23. Hernandez, G., Blumenthal, D. K., Kennedy, M. A., Unkefer, C. J., and Trewhella, J. (1999) *Biochemistry 38*, 6911-7.
24. Tung, C. S., Wall, M. E., Gallagher, S. C., and Trewhella, J. (2000) *Protein Sci 9*, 1312-26.

25. Tripet, B., De Crescenzo, G., Grothe, S., O'Connor-McCourt, M., and Hodges, R. (2002) *J Mol Biol 323*, 345.
26. Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) *J Mol Biol 271*, 728-50.
27. McKay, R. T., Saltibus, L. F., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 12731-12738.
28. McKay, R. T., Pearlstone, J. R., Corson, D. C., Gagne, S. M., Smillie, L. B., and Sykes, B. D. (1998) *Biochemistry 37*, 12419-30.
29. Gagne, S. M., Li, M. X., McKay, R. T., and Sykes, B. D. (1998) *Biochem Cell Biol 76*, 302-12.
30. Gagné, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry 36*, 4386-4392.
31. Gagné, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) *Prot. Sci. 3*, 1961-1974.
32. Gagné, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) *Nature Struct. Biol. 2*, 784-789.
33. Putkey, J. A., Sweeney, H. L., and Campbell, S. T. (1989) *J. Biol. Chem. 264*, 12370-8.
34. Putkey, J. A., Liu, W., and Sweeny, H. L. (1991) *J. Biol. Chem. 266*, 14881-14884.
35. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
36. Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) *Proc Natl Acad Sci U S A.*
37. Ferguson, R. E., Sun, Y. B., Mercier, P., Brack, A. S., Sykes, B. D., Corrie, J. E., Trentham, D. R., and Irving, M. (2003) *Mol Cell 11*, 865-74.
38. Qiu, F., Lakey, A., Agianian, B., Hutchings, A., Butcher, G. W., Labeit, S., Leonard, K., and Bullard, B. (2003) *Biochem J 371*, 811-21.
39. Wishart, D. S., Boyko, R. F., Willard, L., Richards, F. M., and Sykes, B. D. (1994) *Comput Appl Biosci 10*, 121-132.

APPENDIX A

Establishment of a high yield recombinant peptide expression system

Appendix A-1. Codon usage in *Escherichia coli* K12

Taken from NCBI website:
http://www.molbiol.ox.ac.uk/~cocallag/refdata_html/codonusagetable.shtml

Escherichia coli K12 [gbbct]: 5089 CDS's (1608122 codons)

fields: [triplet] [frequency: **per thousand**] ([number])

```
UUU F 0.57 22.4 ( 35982)  UCU S 0.15  8.5 ( 13687)  UAU Y 0.57 16.3 ( 26266)  UGU C 0.45  5.2 (  8340)
UUC F 0.43 16.6 ( 26678)  UCC S 0.15  8.6 ( 13849)  UAC Y 0.43 12.3 ( 19728)  UGC C 0.55  6.4 ( 10347)
UUA L 0.13 13.9 ( 22376)  UCA S 0.12  7.2 ( 11511)  UAA * 0.64  2.0 (  3246)  UGA * 0.29  0.9 (  1468)
UUG L 0.13 13.7 ( 22070)  UCG S 0.15  8.9 ( 14379)  UAG * 0.07  0.2 (   378)  UGG W 1.00 15.3 ( 24615)

CUU L 0.10 11.0 ( 17754)  CCU P 0.16  7.1 ( 11340)  CAU H 0.57 12.9 ( 20728)  CGU R 0.38 21.0 ( 33694)
CUC L 0.10 11.0 ( 17723)  CCC P 0.13  5.5 (  8915)  CAC H 0.43  9.7 ( 15595)  CGC R 0.40 22.0 ( 35306)
CUA L 0.04  3.9 (  6212)  CCA P 0.19  8.5 ( 13707)  CAA Q 0.35 15.4 ( 24835)  CGA R 0.06  3.6 (  5716)
CUG L 0.50 52.7 ( 84673)  CCG P 0.52 23.2 ( 37328)  CAG Q 0.65 28.8 ( 46319)  CGG R 0.10  5.4 (  8684)

AUU I 0.51 30.4 ( 48818)  ACU T 0.17  9.0 ( 14397)  AAU N 0.45 17.7 ( 28465)  AGU S 0.15  8.8 ( 14092)
AUC I 0.42 25.0 ( 40176)  ACC T 0.43 23.4 ( 37624)  AAC N 0.55 21.7 ( 34912)  AGC S 0.28 16.1 ( 25843)
AUA I 0.07  4.3 (  6962)  ACA T 0.13  7.1 ( 11366)  AAA K 0.77 33.6 ( 54097)  AGA R 0.04  2.1 (  3337)
AUG M 1.00 27.7 ( 44614)  ACG T 0.27 14.4 ( 23124)  AAG K 0.23 10.2 ( 16401)  AGG R 0.02  1.2 (  1987)

GUU V 0.26 18.4 ( 29569)  GCU A 0.16 15.4 ( 24719)  GAU D 0.63 32.2 ( 51852)  GGU G 0.34 24.9 ( 40019)
GUC V 0.22 15.2 ( 24477)  GCC A 0.27 25.5 ( 40993)  GAC D 0.37 19.0 ( 30627)  GGC G 0.40 29.4 ( 47309)
GUA V 0.15 10.9 ( 17508)  GCA A 0.21 20.3 ( 32666)  GAA E 0.69 39.5 ( 63517)  GGA G 0.11  7.9 ( 12776)
GUG V 0.37 26.2 ( 42212)  GCG A 0.35 33.6 ( 53988)  GAG E 0.31 17.7 ( 28522)  GGG G 0.15 11.0 ( 17704)
```

Coding GC 51.80% 1st letter GC 58.89% 2nd letter GC 40.72% 3rd letter GC 55.79%
Note: first figure is proportional use for that amino acid.

Genetic code 11: Eubacterial

```
Escherichia coli
AmAcid  Codon      Number      /1000     Fraction   ..

Gly     GGG     17704.00      11.01       0.15
Gly     GGA     12776.00       7.94       0.11
Gly     GGT     40019.00      24.89       0.34
Gly     GGC     47309.00      29.42       0.40

Glu     GAG     28522.00      17.74       0.31
Glu     GAA     63517.00      39.50       0.69
Asp     GAT     51852.00      32.24       0.63
Asp     GAC     30627.00      19.05       0.37

Val     GTG     42212.00      26.25       0.37
Val     GTA     17508.00      10.89       0.15
Val     GTT     29569.00      18.39       0.26
Val     GTC     24477.00      15.22       0.22
```

Ala	GCG	53988.00	33.57	0.35
Ala	GCA	32666.00	20.31	0.21
Ala	GCT	24719.00	15.37	0.16
Ala	GCC	40993.00	25.49	0.27
Arg	AGG	1987.00	1.24	0.02
Arg	AGA	3337.00	2.08	0.04
Ser	AGT	14092.00	8.76	0.15
Ser	AGC	25843.00	16.07	0.28
Lys	AAG	16401.00	10.20	0.23
Lys	AAA	54097.00	33.64	0.77
Asn	AAT	28465.00	17.70	0.45
Asn	AAC	34912.00	21.71	0.55
Met	ATG	44614.00	27.74	1.00
Ile	ATA	6962.00	4.33	0.07
Ile	ATT	48818.00	30.36	0.51
Ile	ATC	40176.00	24.98	0.42
Thr	ACG	23124.00	14.38	0.27
Thr	ACA	11366.00	7.07	0.13
Thr	ACT	14397.00	8.95	0.17
Thr	ACC	37624.00	23.40	0.43
Trp	TGG	24615.00	15.31	1.00
End	TGA	1468.00	0.91	0.29
Cys	TGT	8340.00	5.19	0.45
Cys	TGC	10347.00	6.43	0.55
End	TAG	378.00	0.24	0.07
End	TAA	3246.00	2.02	0.64
Tyr	TAT	26266.00	16.33	0.57
Tyr	TAC	19728.00	12.27	0.43
Leu	TTG	22070.00	13.72	0.13
Leu	TTA	22376.00	13.91	0.13
Phe	TTT	35982.00	22.38	0.57
Phe	TTC	26678.00	16.59	0.43
Ser	TCG	14379.00	8.94	0.15
Ser	TCA	11511.00	7.16	0.12
Ser	TCT	13687.00	8.51	0.15
Ser	TCC	13849.00	8.61	0.15
Arg	CGG	8684.00	5.40	0.10
Arg	CGA	5716.00	3.55	0.06
Arg	CGT	33694.00	20.95	0.38

Arg	CGC	35306.00	21.95	0.40
Gln	CAG	46319.00	28.80	0.65
Gln	CAA	24835.00	15.44	0.35
His	CAT	20728.00	12.89	0.57
His	CAC	15595.00	9.70	0.43
Leu	CTG	84673.00	52.65	0.50
Leu	CTA	6212.00	3.86	0.04
Leu	CTT	17754.00	11.04	0.10
Leu	CTC	17723.00	11.02	0.10
Pro	CCG	37328.00	23.21	0.52
Pro	CCA	13707.00	8.52	0.19
Pro	CCT	11340.00	7.05	0.16
Pro	CCC	8915.00	5.54	0.13

Appendix A-2. Sequencing results of GEV1 peptide expression vectors

*Note – *All sequences have been sequenced with the T7-terminator primer (5'-TATGCTAGTTATTGCTCAG-3') and are listed in a 5' to 3' fashion (reverse/compliment).*

A. GEV1-clp-RR-20

```
5'...NGCNTCCACCAGCGTTCCTTTCNNGGCTTTGTTAGCAGCCGGATCTCAGTGGTG
GTGGTGGTGGTGCTCGAGCCCGGGCCCGCGGGGCATGCGCACGCGGCGCAGGGTCGGGC
GTTTAAATTTGCCGCGCAGATCAAAAATTTTCTGGGTCATGCTAGCCGGACCGCCCGGT
TCGGTTACCGTGAAGGTTTTGGTAGCGTCGTCGTAGGTCCATTCACCGTCAACACCGTT
GTCGTTAGCGTACTGTTTGAAAACTTTTTCCGCGGTAGCAGCGTCAACAGCTTCGGTGG
TGGTTTCACCTTTCAGGGTTTTACCGTTCAGAGCAAGCTTGTACTGCATATGTATATCT
CCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCC
CCTATAGTGAGTCGTATTAATTTCGCGGATCGAGATCTCGATCCTCTACGCCGGACGC
ATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACAT
CACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGG
GTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCA
TTCCTNGCNGCGGCGGTGCTCAACGGCCTCAACCTACTACTT...3'
```

B. GEV1-clp-GR-20

```
5'...TTGGAGGCGGCCTTTCCCCACTACGCGTCCTTTCGAGGCTTTGTTACGCACGCC
GGATCTCAGTGGTGGTGGTGGTGCTCGAGCCCGGGCATGCGCACGCGGCCCAGGGT
CGGGCGTTTAAATTTGCCGCGCAGATCAAAAATTTTCTGGGTCATGCTAGCCGGACCGC
CCGGTTCGGTTACCGTGAAGGTTTTGGTAGCGTCGTCGTAGGTCCATTCACCGTCAACA
CCGTTGTCGTTAGCGTACTGTTTGAAAACTTTTCTCCGCGGTAGCAGCGTCAACAGCTT
CGGTGGTGGTTTCACCTTTCAGGGTTTTACCGTTCAGAGCAAGCTTGTACTGCATATGT
ATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCAC
```

AATTCCCCCTATAGTGAGTCCGTATTAATTTCGCGGATCGAGATCTCGATCCTCTACG
CCGGACGCATCGTGGCCGGCATCACCGGCGCCACCACGCGTGCGGTTGCTGGCGCCTAT
ATCGCCGACATCACCGATGGGGAACCATCGGGCTCGCCACTTCCGGCTCATGAGCGCTT
GTCTCGGCGTGGGTATCGTCGCCCGCCCCGTGGCCACGGACTGTTGGGCGCCATCTCC
TTGCATGCACCATCCCTTGCGGCGGCGGTGCTAACCGGCCTCCACCTACTACTGGGCTG
CTTCCTACTGCACGACCCCCCTCACCGCACAACACTCAAAATCCCCGGAACCCATCGAA
ATAGCAACACAACCCTCTCCCCGGTACTGGCATTGATAACCGCCCCGCAAAAACACCTT
CATCTAAGGCCTGGCTCACATTCACAAACCCCTCACCTTTAATACCAAGTTCCCACAAC
ATTGACCGCGTCTTACTCTATATACACATTCCAATG...3′

C. GEV1-clp-RG-20

5′...CTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTG
GTGCTCGAGCCCGGGCATGCGCACGCCGCGCAGGGTCGGGCGTTTAAATTTGCCGCGCA
GATCAAAAATTTTCTGGGTCATGCTAGCCGGACCGCCCGGTTCGGTTACCGTGAAGGTT
TTGGTAGCGTCGTCGTAGGTCCATTCACCGTCAACACCGTTGTCGTTAGCGTACTGTTT
GAAAACTTTTTCCGCGGTAGCAGCGTCAACAGCTTCGGTGGTGGTTTCACCTTTCAGGG
TTTTACCGTTCAGAGCAAGCTTGTACTGCATATGTATATCTCCTTCTTAAAGTTAAACA
AAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATT
AATTTCGCGGGATCGAGATCTCGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACC
GGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCG
GGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCANGCCCCG
TGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTG
CTCAACNGCCTCAACCTACTACTGGGCTGCTTNCTAATGCAAGAGTCGCATAAGGGAGA
GCGTNGAGATCCGGACACCATCGAATGGCGCAAAACNTTCGCNGTATGCATGATGCNCC
CGAGANAGTCATCAGNNNNAAN...3′

D. GEV1-TRTK-12

5′...GTNGANNGCTAGCCANCTCAGCTTCCTTTCGNGGCTTTGTTAGCAGCCGGATCT
CAGTGGTGGTGGTGGTGGTGCTCGAGCCCGGGCATGCTCAGAATTTTGTTCCAATCAAT
TTTGGTGCGGGTCATGCTAGCCGGACCGCCCGGTTCGGTTACCGTGAAGGTTTTGGTAG
CGTCGTCGTAGGTCCATTCACCGTCAACACCGTTGTCGTTAGCGTACTGTTTGAAAACT
TTTTCCGCGGTAGCAGCGTCAACAGCTTCGGTGGTGGTTTCACCTTTCAGGGTTTTACC
GTTCAGAGCAAGCTTGTACTGCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTAT
TTCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCG
CGGGATCGAGATCTCGATCCTCTACGCCCGGACGCATCGTGGCCGGCATCACCGGCGCC
ACAGGTGCGGTTGCTGGCGCCTATATCGCNGACATCACCGATGGGGAAGATCNGGCTCG
CCACTNCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCNNNNCCCGTGGCCG
GGGGGACTGTTGGGGCGCNATCTCCCTTGCATGCAACCANTNCCCTTGCCGGCGGGCGG
TTGCTTCAAACGGGCCTCAAACCNTACTTANNNGGGGCTTGNNTNCCCTANNGGCANGG
AANTTTCCATTTAANGGGAGANCCTTTNNAAAATCCCNGGACACCCATNTGAANTTGGG
CCCAAAAACCNNNNNNCGGNTATGGNATTTATNANCCCCNNGCNANAAAAATTTAATTT
GCGGGTTGGNTNAANTTNNAACACT...3′

E. GEV1-EDQL-26

5'...GCACTAGCCNANCNTCAGCNTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGT
GGTGGTGGTGGTGGTGCTCGAGCCATGGCATCTGCAGATCAAATTTCTCAGCCTCCAGA
TTATAAATGCTCTGCCACAGCTCTTTAGCTTTCTCGCGCAGCTGATCCTCCATGCTAGC
CGGACCGCCCGGTTCGGTTACCGTGAAGGTTTTGGTAGCGTCGTCGTAGGTCCATTCAC
CGTCAACACCGTTGTCGTTAGCGTACTGTTTGAAAACTTTTTCCGCGGTAGCAGCGTCA
ACAGCTTCGGTGGTGGTTTCACCTTTCAGGGTTTTACCGTTCAGAGCAAGCTTGTACTG
CATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATC
CGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCTCGATCCT
CTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACANGTGCNGTTGCTGGCGCCT
ATATCGCCGACATCACCGATNGGGAAGATCNGGCTCGCCACTTCNGGCTCATGANCNCT
TGTTTCGGCGTNGGNATGGTGGCACGCCCCNNGGNCGGGGGACTGTGGGGCCCCATCTN
CCTGCATGCCCAATTNNTGCCGGCGCGGGGCCCAACCGNCNTNNACCACTACTGGGCTG
CTTCCTATTGCAGGAAGTNCATANNGGAAACCNTNNANATCCGGNACCCCATTAANTGG
NGCCCAANCCTTTCCCGNNNTGCANTNANTNNCCCCCGAAAAAANTTTNATTTNNGGNN
NNTGANNTTTAACNNNAT...3'

F. GEV1-FLQS-26

5'...GGCAGCATCCACCTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTG
GTGGTGGTGGTGGTGCTCGAGCCCGGGCATGCGCAGCGGCAGAAAATAGCCCGCATCCA
GAATAATCGGCGGCAGCAGAAACAGAAAAAACACATCGCTCTGCAGAAACATGCTAGCC
GGACCGCCCGGTTCGGTTACCGTGAAGGTTTTGGTAGCGTCGTCGTAGGTCCATTCACC
GTCAACACCGTTGTCGTTAGCGTACTGTTTGAAAACTTTTTCCGCGGTAGCAGCGTCAA
CAGCTTCGGTGGTGGTTTCACCTTTCAGGGTTTTACCGTTCAGAGCAAGCTTGTACTGC
ATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCC
GCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCTCGATCCTC
TACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTA
TATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTT
GTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTGGGCGCCATCTCC
TTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCACCTACTACTGGGCTGC
TTCTAATGCAGGAGTCGCATAAGGGAAAGCGTTAAGATCCCGACACCATCGAATGCCGC
AAAACCTTCCCCGTATTGCTTTATTCCCCCTGAAAAAGATTAATTACAGGGTGTGG...
3'

Appendix A-3. **Sequencing results of pET3a cTnI and cTnT expression vectors**

*Note – *All sequences have been sequenced with the T7-terminator primer (5'-TATGCTAGTTATTGCTCAG-3') and are listed in a 5' to 3' fashion (reverse/compliment).*

A. pET3a-cTnI(34-163)

5'...CAATGGGTGAAATGGCAAAGGAAAATGGGGGCCCCAAACCAATTCCCCCCGGGG
CCACGGGGGCCCCTGTGCCAACCAATAACCCCAAGGGCCCGGAAACCAAGCGGGTTTAA
TGAGGCCCCCGAAAGTTGGCGGAGGCCCCGGATCTTTCCCCCCAATCGGGTAAATGGTC

GGGCGAATATAGGGGCGCCAAGCAACCCGCACCCCTGTGGCGGCCGGGTGATGCCCGGC
CCACGATGCGTCCGGGCGTAGAGGGATCGAGATTCTCGATCCCCGCGAAATTAATACGG
ACTCACTATAGGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAA
GAAGGAGATATACATATGGCCAAGAAAAAATCTAAGATCTCCGCCTCGGAGAAAATTGC
AGCTGAAGACTCTGCTGCTGCAGATTGCAAAGCAAGAGCTGGAGCGAGAGGCGGAGGAG
CGGCGCGGAGAGAAGGGGCGCGCTCTGAGCACCCGCAGCCAGCCGCTGGAGTTGACCGG
GCTGGGCTTCGCGGAGCTGCAGGACTTGAGCCGCCAGCTCCACGCCCGTGTGGACAAGG
TGGATGAAGAGAGATACGACATAGAGGCAAAAGTCACCAAGAACATCACGGAGATTGCA
GATCTGACTCAGAAGATCTTTGACCTTCGAGGCAAGTTTAAGCGGCCCACCCTGCGGAG
AGTGAGGATCTCTGCAGATGCCATGATGCAGGCGCTGTCTGGGGGCCCGGGCTTGAGGA
TCCGGACTGCTAACAAAGCCTCGAAAGGAACGCTCGATGTTGGCTAGCTTCAACCG...
3′

B. pET3a-cTnT(197-298)

5′...GATAAAGTTTCCCGTAAGGAAACTCCGGGCATTAGGGAAGCAGGCCCCAAGGTA
GTAAGGCTGAGCCCTTTTTTTTTCTTTTCTTTTTGTCGTGGTGGGGGGAGGGGAGAGT
TTTTTTTCTTCCCTCTCTCATCTCCCCCCGCCACGGGGCCTGCCACCATACCCACGCCG
AAACCAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGC
GATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGG
CGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGAGACCAC
AACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATCCAGAA
GCAGGCCCAGACAGAGCGGAAAAGTGGGAAGAGGCAGACTGAGCGGAAAAGAAGAAGA
AGATTCTGGCTGAGAGGAGGAAGGTGCTGGCCATTGACCACCTGAATGAAGATCAGCTG
AGGGAGAAGGCCAAGGAGCTGTGGCAGAGCATCTATAACTTGGAGGCAGAGAAGTTCGA
CCTGCAGGAGAAGTTCAAGCAGCAGAAATATGAGATCAATGTTCTCCGAAACAGGATCA
ACGATAACCAGAAAGTCTCCAAGACCCGCGGGAAGGCTAAAGTCACCGGGCGCTGGAAA
...3′

Appendix A-4. MALDI-TOF mass spectrometry results of HPLC purified peptide

These are the MALDI-TOF mass spectrometry results for the CNBr cleaved fusion protein GB1_RR-20_His. The HPLC chromatogram purification plot can be viewed in Chapter II, Figure II-8 of this thesis.

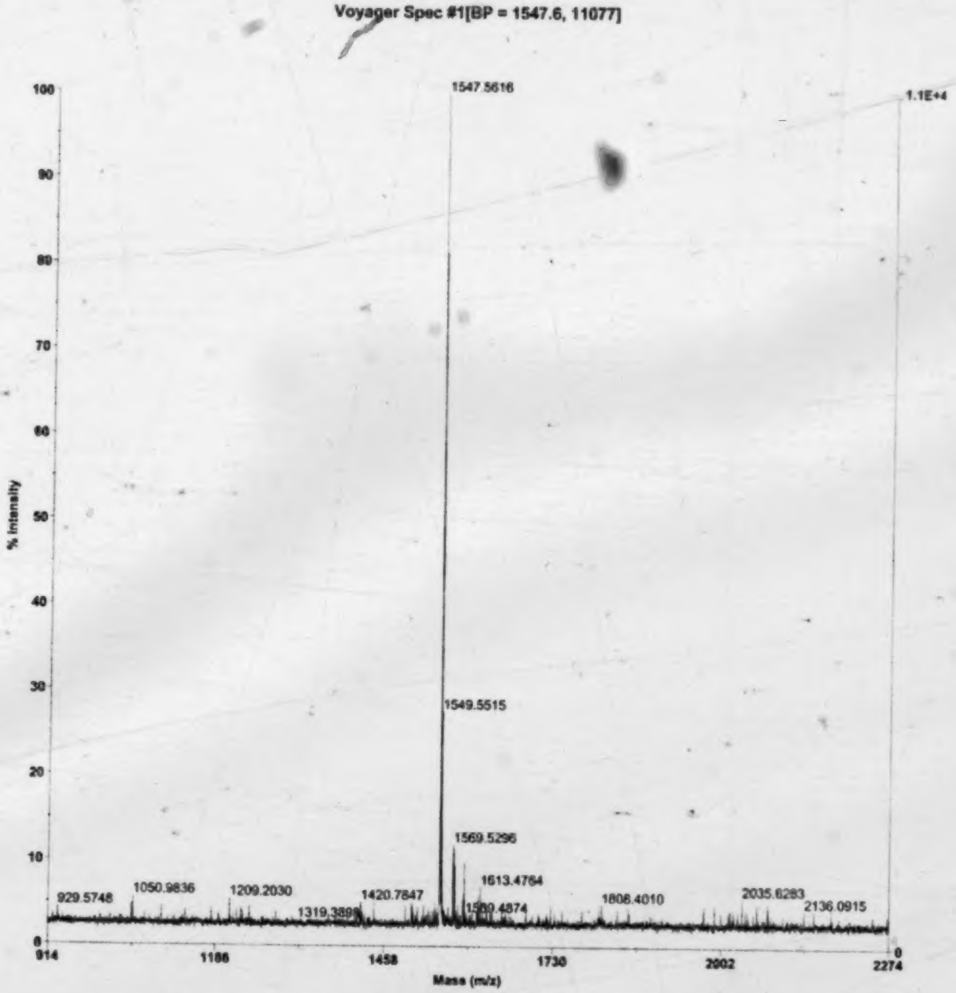


Figure A-1. **His-tag fragment**. HPLC fraction time 10 minutes, the His-tag fragment has a predicted molecular weight of 1547.7 Da and a MALDI-TOF verified mass of 1547.6 Da.

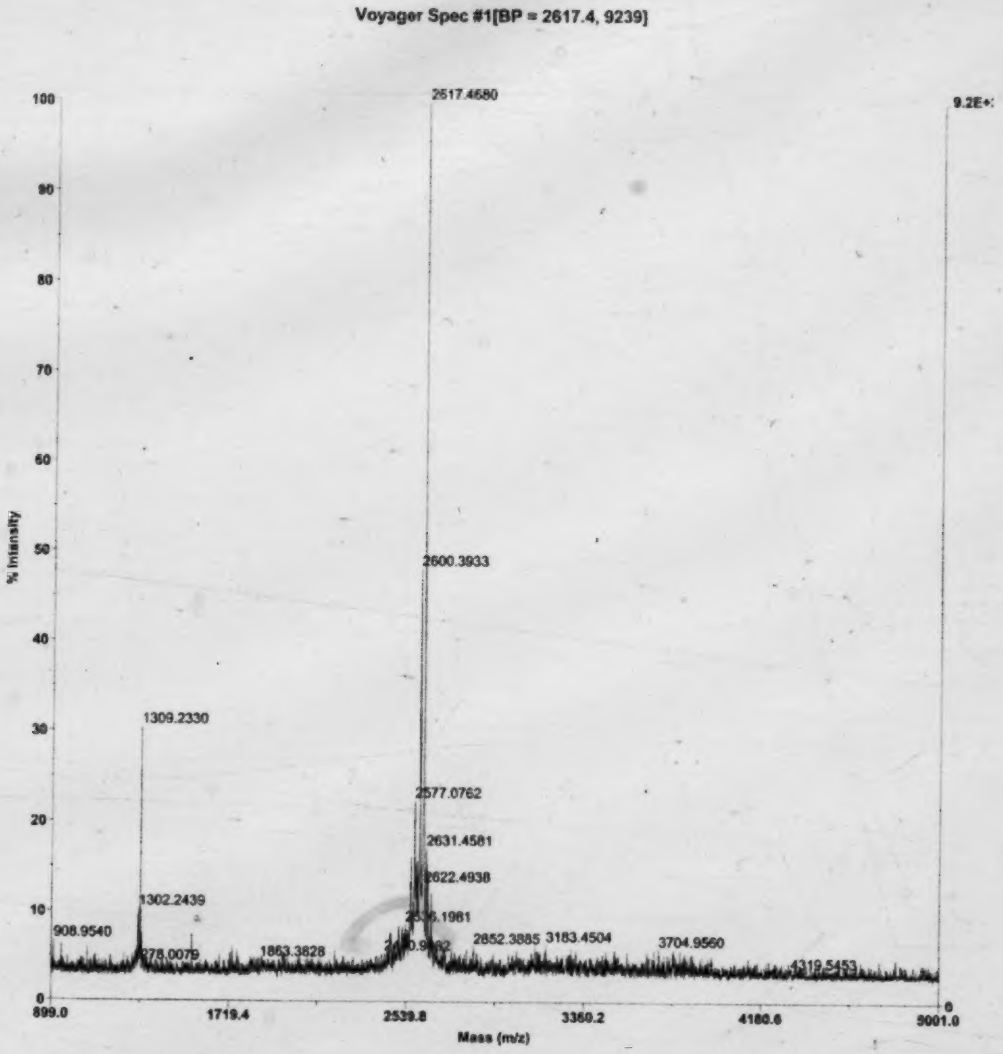


Figure A-2. **cIp-RR-20 peptide fragment**. HPLC fraction time 23 minutes, this cIp-RR-20 fragment contains homoserine and has a predicted molecular weight of 2618.2 Da and a MALDI-TOF verified mass of 2617.5 Da.

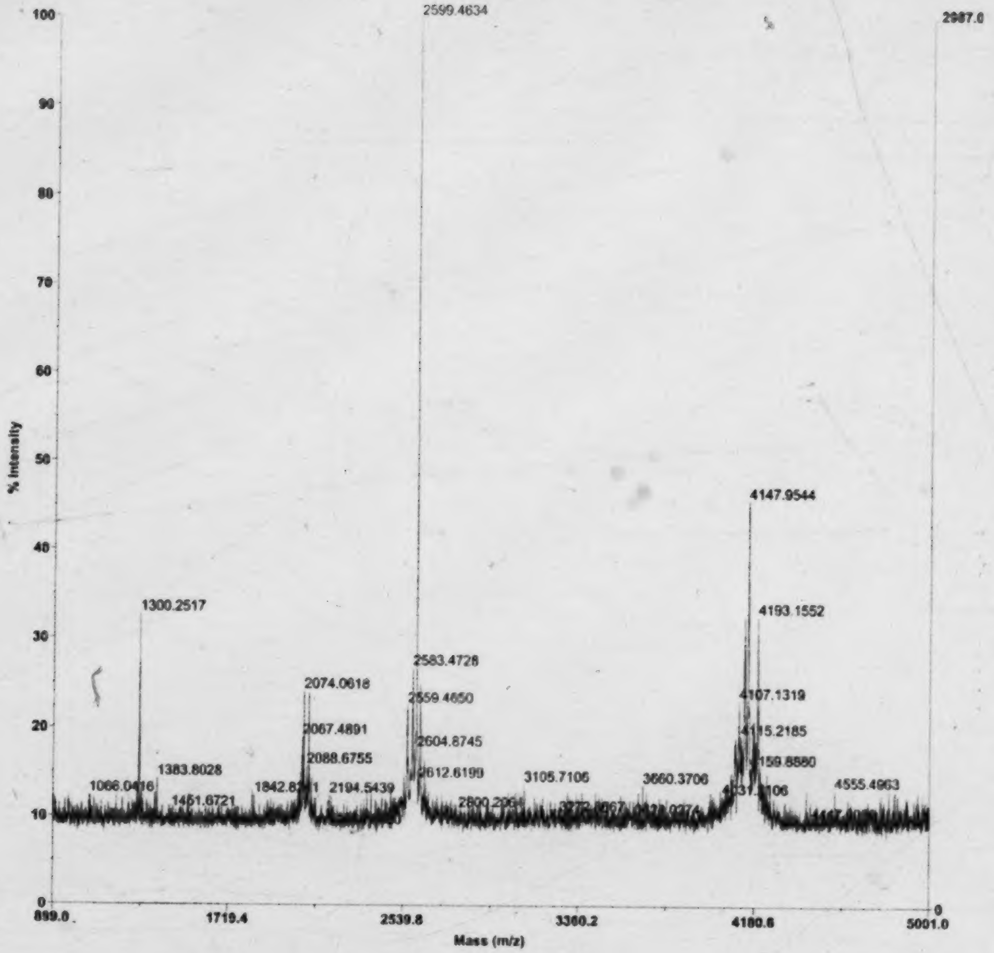


Figure A-3. **cIp-RR-20 peptide fragment**. HPLC fraction time 23.5 minutes, this cIp-RR-20 fragment contains homoserine lactone and has a predicted molecular weight of 2600.2 Da and a MALDI-TOF verified mass of 2599.5 Da.

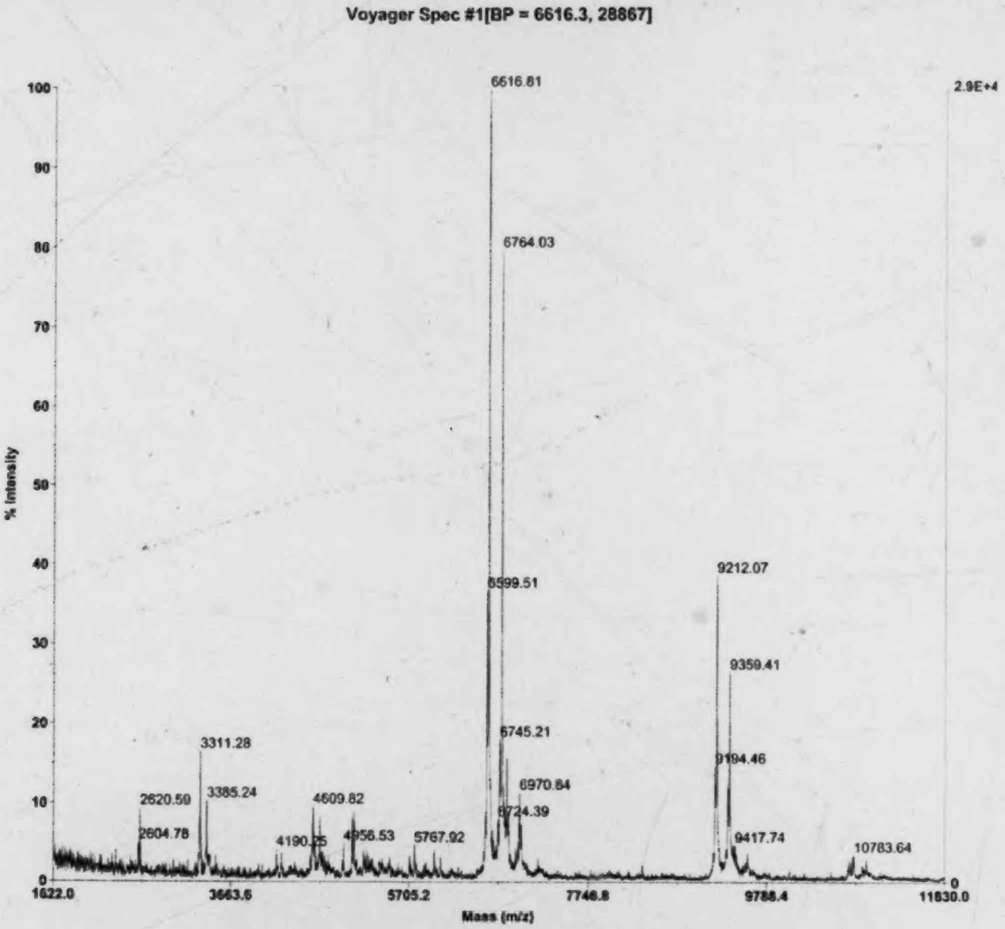


Figure A-4. **GB1 peptide fragment**. HPLC fraction time 31 minutes, this GB-1 fragment contains homoserine and has a predicted molecular weight of 6624.4 Da and a MALDI-TOF verified mass of 6616.2 Da.

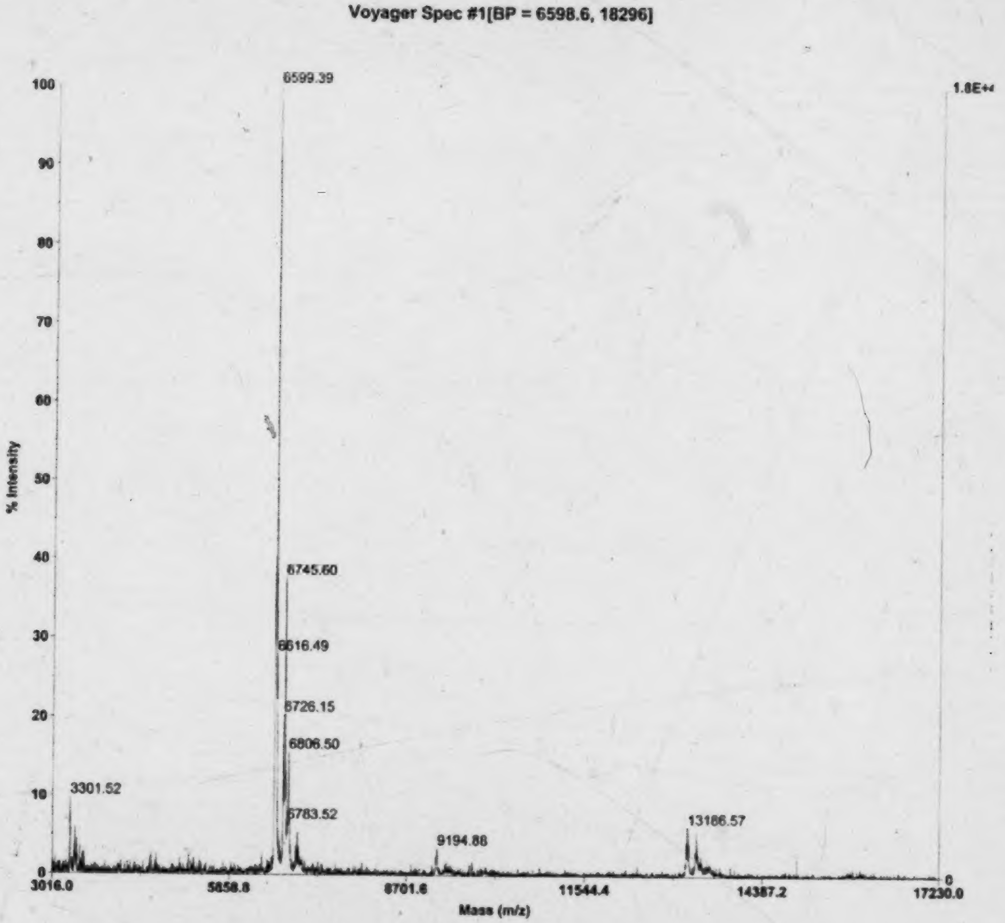


Figure A-5. **GB1 peptide fragment**. HPLC fraction time 31 minutes, this GB-1 fragment contains homoserine lactone and has a predicted molecular weight of 6606.4 Da and a MALDI-TOF verified mass of 6599.4 Da.

APPENDIX B

Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na^+/H^+ exchanger

This appendix contains information regarding GB1 fusion peptide FLQS-26, a transmembrane spanning segment 4 (TM IV) of the human sodium-proton exchanger isoform 1 (hNHE1), described within Chapter II of this thesis. In addition, it contains a large amount of functional information on the hNHE1. All functional data was performed solely by Emily Slepkov and the Fliegel laboratory in the Department of Biochemistry at the University of Alberta. NMR structural data and analysis was primarily performed by Dr. Jan Rainey and myself in the Sykes laboratory in the Department of Biochemistry at the University of Alberta. This work was recently published in May 2005 in the *Journal of Biological Chemistry* [*J Biol Chem.* 2005 May 6 **280(18)**:17863-72]. All authors involved in this body of work are Emily R. Slepkov, Jan K. Rainey, Xiuju Li, Yongsheng Liu, Florence J. Cheng, Darrin A. Lindhout, Brian D. Sykes and Larry Fliegel. This appendix is supplied solely as a supplement to the ubiquitous applications of the GB-1 expression system described in Chapter II, and is an exact replica of the JBC manuscript.

OVERVIEW

The Na^+/H^+ exchanger isoform 1 is a ubiquitously expressed integral membrane protein that regulates intracellular pH in mammals. We characterized structural and functional aspects of the critical transmembrane (TM) segment IV. Each residue was mutated to cysteine in cysteineless NHE1. TM IV was exquisitely sensitive to mutation with 10 of 23 mutations causing greatly reduced expression and/or activity. The Phe161Cys mutant was inhibited by treatment with the water soluble sulfhydryl reactive compounds MTSET ([2-(Trimethylammonium)Ethyl]Methanethiosulfonate) and MTSES ([2-Sulfonatoethyl]Methanethiosulfonate), suggesting it is a pore lining residue. The structure of purified TM IV peptide was determined using high-resolution NMR in a CD_3OH:$CDCl_3$:H_2O mixture and in DMSO. In CD_3OH:$CDCl_3$:H_2O, TM IV was structured, but not as a canonical α-helix. Residues D159-L162 were a series of β-turns,

residues L165-P168 showed an extended structure, while residues I169-F176 were helical in character. These three structured regions rotated quite freely with respect to the others. In DMSO, the structure was much less defined. Our results demonstrate that TM IV is an unusually structured transmembrane segment that is exquisitely sensitive to mutagenesis and that Phe161 is a pore-lining residue.

ABBREVIATIONS

BMRB, BioMagResBank; cNHE1, cysteineless NHE1; DAPI, 4',6-Diamidino-2-phenyindole, diacetate; DMSO: dimethylsulfoxide; DQF-COSY, double-quantum filtered correlation spectroscopy; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulfonate; HA, Hemagglutinin; HNHA, ^{15}N-edited $^{1}H^{N}$-$^{1}H^{\alpha}$ correlation spectroscopy; Hsl, homoserine lactone; HSQC, heteronuclear single quantum coherence spectroscopy; MALDI-MS, matrix assisted laser-desorption ionization mass spectrometry; MTSES ([2-Sulfonatoethyl]-Methanethiosulfonate); MTSET, [2-(Trimethylammonium)-Ethyl]-Methanethiosulfonate; NHE1-8, Na^+/H^+ exchanger isoforms 1 to 8; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; PBS, Phosphate Buffered Saline; PDB, Protein Data Bank; RMSD, root-mean-square deviation; TFA, trifluoroacetic acid; TM, Transmembrane segment; TOCSY, total correlation spectroscopy

INTRODUCTION

The mammalian Na^+/H^+ exchanger isoform 1 (NHE1) is a ubiquitously expressed integral membrane protein that mediates the removal of one intracellular proton in exchange for one extracellular sodium ion (1). NHE1 thereby protects cells from intracellular acidification (*2*, *3*), and stimulation of its activity promotes cell growth, differentiation (2), and regulates sodium fluxes and cell volume after osmotic shrinkage (*2*, *3*). The Na^+/H^+ exchanger also plays an important role in the damage that occurs to the human myocardium during ischemia and reperfusion, and it has been shown that inhibition of the exchanger has beneficial effects on the myocardium under these conditions(4). Amiloride and its derivatives are inhibitors of the NHE1 isoform of the Na$^+$/H$^+$ exchanger and a new generation of Na$^+$/H$^+$ exchanger inhibitors is being developed for clinical treatment of heart disease (5).

Although the activity of NHE1 has been extensively examined in many tissues, only recently is information starting to be elucidated on how this antiporter actually binds and transports Na$^+$ ions and protons. NHE1 is composed of two domains: an N-terminal membrane domain of approximately 500 amino acids and a C-terminal regulatory domain of about 315 amino acids (*1, 4*) (Fig. 1). The N-terminal membrane domain is responsible for ion movement and it is reported to have 12 transmembrane (TM) segments and 3 membrane associated segments (6). Transmembrane segment four (TM IV; residues 155-177) has been implicated in the ion transport and inhibitor binding properties of NHE1 (7-9). The sequence of human TM IV of NHE1 is ^{155}FLQSDVFFLFLLPPIILDAGYFL^{177}. The underlined residues have been shown to affect Na$^+$ affinity or the inhibitor resistance of mammalian NHE1 (7-9). Recently, we have shown that prolines 167 and 168 are critical to NHE1 function, targeting and expression (10). These data provide a strong case for the importance of many amino acid residues of TM IV in the ion binding, structure and transport properties of NHE1.

In this communication, we examine both structural and functional aspects of TM IV of the NHE1 isoform of the Na$^+$/H$^+$ exchanger. We use cysteine-scanning mutagenesis to characterize which amino acids are important in function and are likely pore lining residues of TM IV. To study the structure of TM IV the segment was expressed and purified and we compared its structure in two different membrane mimetic solvent conditions. This provides insight into structured regions, in our case, without fixing the segment as a whole into a single conformation. The use of solution conditions having a low dielectric constant to mimic a membrane environment has become quite common for structural studies of transmembrane proteins or peptides and has provided structures of isolated protein segments consistent with their structures in the full protein (11-13). Our study demonstrates that while TM IV is structured, only a 4-6 residue stretch of the segment is helical. In addition we identify Phe 161 as a pore lining residue in NHE1.

EXPERIMENTAL PROCEDURES

B.1 Materials

^{15}N labeled ammonium sulfate and all deuterated NMR solvents were from Cambridge Isotope Laboratories, [Andover, MA, USA]. PWO DNA polymerase was from Roche Molecular Biochemicals, Mannheim, Germany. LIPOFECTAMINE™ 2000 Reagent was from Invitrogen Life Technologies, Carlsbad, CA, USA. MTSET and MTSES were from Toronto Research Chemicals, [Toronto, Ont., Canada].

B.2 Site-directed mutagenesis

Mutations in TM IV were made to an expression plasmid containing a hemagglutinin (HA) tagged human NHE1 isoform of the Na^+/H^+ exchanger. The plasmid pYN4-C contains the cDNA of the entire coding of NHE1 with all ten native cysteine residues mutated to serine as described earlier(10). Each residue of TM IV was individually mutated to cysteine using the cysteineless pYN4-C as a template. Site-directed mutagenesis was performed using amplification with PWO DNA polymerase followed by use of the Stratagene (La Jolla, CA, USA) QuikChange™ site directed mutagenesis kit as recommended by the manufacturer. Mutations were designed to create a new restriction enzyme site for use in screening transformants. DNA sequencing confirmed the accuracy of the mutations.

B.3 Cell culture and stable transfection

AP-1 cells were used to examine Na^+/H^+ exchanger expression and activity. These cells lack an endogenous Na^+/H^+ exchanger. Stable cell lines were made of all mutants by transfection with LIPOFECTAMINE™ 2000 Reagent according to the manufacturer's protocol as described earlier (10). Transfected cells were selected using 800 μg/ml geneticin (G418) and stable cell lines for experiments were regularly re-established from frozen stocks at passage numbers between 5-15.

B.4 SDS-PAGE and immunoblotting

To confirm NHE1 expression immunoblot analysis was used on samples from total cell lysates of AP-1 cells. Cell lysates were made as describe earlier(10). For Western blot analysis equal amounts of each sample (100 μg total protein) was resolved on 10% SDS/polyacrylamide gels. The gel was transferred onto a nitrocellulose membrane and immunostained using anti-HA monoclonal antibody (Boehringer Mannheim, Laval, Que., Canada) and peroxidase-conjugated goat anti-mouse antibody (Bio/Can, Mississauga, ON, Canada). The Amersham enhanced chemiluminescence Western blotting and detection system was used to visualize immunoreactive proteins. Densitometric analysis of X-ray films was carried out using NIH Image 1.63 software (National Institutes of Health, Bethesda, MD, USA).

B.5 Cell surface expression

Cell surface expression was measured essentially as described earlier(10). Briefly, cells were labeled with Sulpho-NHS-SS-Biotin (Pierce Chemical Company, Rockford, IL, USA) and immobilized streptavidin resin was used to remove cell surface labeled Na^+/H^+ exchanger. Equivalent amounts of the total and unbound proteins were analyzed by SDS-PAGE and Western blotting was as described above. The relative amount of NHE1 on the cell surface was calculated by comparing both the 110 kDa and the 95-kDa species of NHE1 in Western blots of the total and unbound fractions. Results are shown as mean ± SE with statistical significance determined by a Mann-Whitney U test.

B.6 Na^+/H^+ exchange activity and sulfhydryl reactive reagent inhibition

Na^+/H^+ exchange activity was measured using a PTI Deltascan spectrofluorometer or a Schimadzu RF 5000 spectrofluorometer. The initial rate of Na^+-induced recovery of cytosolic pH (pH_i) was measured after acute acid load using 2',7-bis (2-carboxyethyl)-5 (6) carboxyfluorescein-AM (BCECF-AM; Molecular Probes Inc., Eugene, OR, USA). Ammonium chloride (50 mM X 3 min.) was used to transiently induce an acid load and the recovery in the presence of 135 mM NaCl was measured as described previously(10). There were no differences in buffering capacities of stable cell lines as indicated by the degree of acidification induced by ammonium chloride. Results are shown as mean ± SE and statistical significance was determined using a Mann-Whitney U test.

To test the effect of MTSET and MTSES on activity of the NHE1 mutants we used the standard Na$^+$/H$^+$ exchanger assay with ammonium chloride induced acidification of the cells. In this case cells were acidified two times as described above. After a first control acidification and recovery, either MTSET or MTSES was added to a final concentration of 10 mM for 10 minutes in Na$^+$-free buffer. The cells were subsequently washed three times in Na$^+$-free buffer prior to the second ammonium chloride induced acidification and recovery. To calculate residual activity the following formula was used

$$\%\ \text{residual activity} = \frac{\text{pH change after (Reagent)}}{\text{pH change w/o (Reagent)}} \times 100\%$$

B.7 Production and purification of TM IV

We produced TM IV as a fusion protein with the immunoglobulin-binding domain of *Streptococcal* protein G (GB1 domain). DNA encoding the residues of human NHE1 amino acids 155 to 180 (^{155}FLQSDVFFLFLLPPIILDAGYFLPLR^{180}) was synthesized synthetically in the core facility of the Department of Biochemistry. Two self-annealing primers were designed to flank the protein sequence with Met residues and to change the codon preference to that favored in *E. coli*. The primers were 5'-CTAGCATGTT TCTGCAGAGC GATGTGTTTT TTCTGTTTCT GCTGCCGCCG ATTATTCTGG ATGCGGGCTA TTTTCTGCCG CTGCGCATGC CCGGGC- 3', and 5'- TCGAGCCCGG GCATGCGCAG CGGCAGAAAA TAGCCCGCAT CCAGAATAAT CGGCGGCAGC AGAAACAGAA AAAACACATC GCTCTGCAGA AACATG-3'. The primers were designed to have overhanging "sticky" *NheI* and *XhoI* ends so as to insert into the GEV-1 vector (*14, 15*). The annealed primers were ligated into the cut vector and transformed in *E. coli* XL1Blue, then later transformed into *E. coli*

(BL21 DE3 pLysS) for protein production. The correct clones were verified by restriction enzyme digests and DNA sequencing. Cultures were induced with IPTG and treated to release the soluble protein as described earlier (14). For routine production of the protein, cultures were started from a single colony and grown in logarithmic stages until 2.0 liters. At OD_{600} of 0.8, cultures were induced with 1 mM IPTG for 5 hours at 37 degrees. Bacterial pellets were resuspended in 40 ml of 1 X PBS with 1% Triton X-100. The fusion protein was constructed such that a 6-His tag flanked the C terminal Met. This was used for purification of the protein via Ni-NTA affinity chromatography as described by the manufacturer (Qiagen). The fusion protein was eluted from the column using 1 M imidazole. After purification the eluted proteins concentrated using an Amicon concentrator and resuspended in 150 mM NaCl, 50 mM Na_2HPO_4. They were then desalted by using a Sephadex G-25 column in 10 mM NH_4HCO_3 pH 8.0 and then were lyophilized to dryness. To cleave the transmembrane segment IV free of GB1, standard techniques were employed. Briefly, a given purified batch of fusion peptide was lyophilized twice out of 0.05% (v:v) trifluoroacetic acid (TFA; 99%, Aldrich, St. Louis, MO) in water, dissolved to ~2 mg/mL in 50% (v:v) TFA in water with excess cyanogen bromide (97%, Aldrich), and allowed to react in darkness for ~16 hours at room temperature. The reaction mixture was quenched with an equal volume of deionized water and subsequently lyophilized. Reverse-phase HPLC was used to separate the various peptide products of the cleavage, and fractions containing the TM IV segment were pooled and lyophilized at least twice prior to use. The correct fragment identity was determined by MALDI-TOF mass spectrometry.

Preliminary experiments produced unlabelled TM IV. Later experiments produced labeled TM IV for more detailed characterization of the peptide structure. For these experiments *E. coli* were grown in a minimal medium containing 87 mM NaH_2PO_4, 34 mM K_2HPO_4, 4 mM $MgSO_4$, 1.8 μM $FeSO_4$, 55.5 mM glucose, pH7.3. The medium was supplemented with 1 gram of $(^{15}NH4)_2SO_4$.

B.8 NMR spectroscopy and structure calculation

One mg samples of unlabelled TM IV peptide were prepared in CD_3OH; trifluoroethanol (TFE):H_2O mixtures; CD_3OH:$CDCl_3$:H_2O (4:4:1 v:v:v); and, DMSO. Note that for CD_3OH:$CDCl_3$:H_2O, screw-cap NMR tubes (535-TR-7, Wilmad Labglass, Buena, NJ) were used to prevent solvent evaporation. Chemical shifts were referenced internally to DSS at 0.5 mM, with indirect referencing employed for ^{15}N (16). One-dimensional ^1H NMR spectra were acquired at 500 MHz on a Varian INOVA spectrometer at a variety of temperatures for each solvent.

Samples for extended analysis were prepared in DMSO and in CD_3OH:$CDCl_3$:H_2O (4:4:1 v:v:v) at ~2 mM peptide and 1.0 mM DSS. 1D ^1H, natural abundance ^{15}N HSQC, TOCSY (60 ms mix; DIPSI spin lock), and NOESY (150 ms mix) experiments were acquired on the NANUC Varian INOVA 800 MHz spectrometer for each sample. DQF-COSY and NOESY spectra at various mixing times were acquired at 600 MHz on either Varian UNITY or INOVA spectrometers. Using uniformly ^{15}N-labelled peptide, 3D ^{15}N edited NOESY-HSQC (150 ms mix), TOCSY-HSQC (57 ms mix, DIPSI spin lock), and HNHA experiments were acquired at 500 MHz on a Varian Inova spectrometer. Spectra were processed either using VNMR (Varian Inc., Palo Alto, CA) or NMRPipe (17); spectral analysis was facilitated with Sparky 3 (T.D. Goddard and D.G. Kneller, University of California, San Francisco).

Structure calculation was carried out in CNS v. 1.1 (18) using NOE contacts derived from the 150 ms mixing time experiments at 800 MHz and ^3J-coupling constants derived from the HNHA experiments at 500 MHz as described by Vaster and Bax (19). Homonuclear NOESY and HNHA peaks were manually picked in Sparky, and volumes were calculated using Gaussian fits, with motion of the peak centre generally allowed; in some cases (< 2%) with the NOESY spectra, Sparky's Gaussian fit algorithm did not find a convergent solution, and a summed signal intensity was used instead over a manually specified region. Initial NOE calibration was carried out empirically from peak volumes to provide a value in the range of 1.8-5.0 Å. These estimates were used to bin each contact into one of strong (1.8-2.8 Å), medium (1.8-3.6 Å), or weak (1.8-5.0 Å). Peaks with an ambiguous assignment between two possible contacts were assigned to the more conservative contact (where possible) as a starting point.

An iterative procedure was employed to refine and prune the NOE contacts using a philosophy similar to that of Wang *et al.* (20). Families of 500-1000 structures were generated by simulated annealing using the default CNS heating and cooling cycles and NOE energetics. NOE violations were analyzed over each ensemble and any violations observed in >~10% of the ensemble structures were examined in detail. If unambiguous, these were typically lengthened; if ambiguous, reassigned or discarded. During this process, restraints were extended to a fourth "very weak" category (1.8-6.0 Å) as necessary (20). After all major NOE violations were resolved, the 3-bond J-coupling values were incorporated directly into the simulated annealing protocol (21) using Karplus coefficients as specified by Vuister and Bax for the HNHA experiment (19) for the more structured CD_3OH:$CDCl_3$:H_2O condition. A further pair of iterations was carried out to optimize the compatibility between ^{3}J-coupling values and NOE restraints. The homoserine lactone residue was incorporated into structure calculations using a protonated residue structure attached to an N-terminal methyl ketone group calculated by MM+ molecular mechanics with bond dipole electrostatics with atomic charges obtained by Restricted Hartree-Fock calculation in HyperChem 3 (Hypercube, Gainesville, FL). XPLO2D 3.2.2 (22) was used to produce CNS topology and parameter files. Structurally convergent segments of the peptide were determined by examining the lowest energy 600 structures out of the final ensemble of 1000 with NMRCORE 1.0 (23); structural features of the 600 member ensemble were analyzed using PROMOTIF-NMR 3.0 (24). The final sets of restraints have been deposited to the PDB (entry 1Y4E) along with the 100 lowest energy structures for the CD_3OH:$CDCl_3$:H_2O solvent condition.

B.9 Coordinates

The coordinates for the structure of hNHE1-TMIV have been deposited in the RCSB Protein Data Bank with the assigned accession code ***1Y4E***.

RESULTS

B.1 Activity and expression of NHE1 mutants

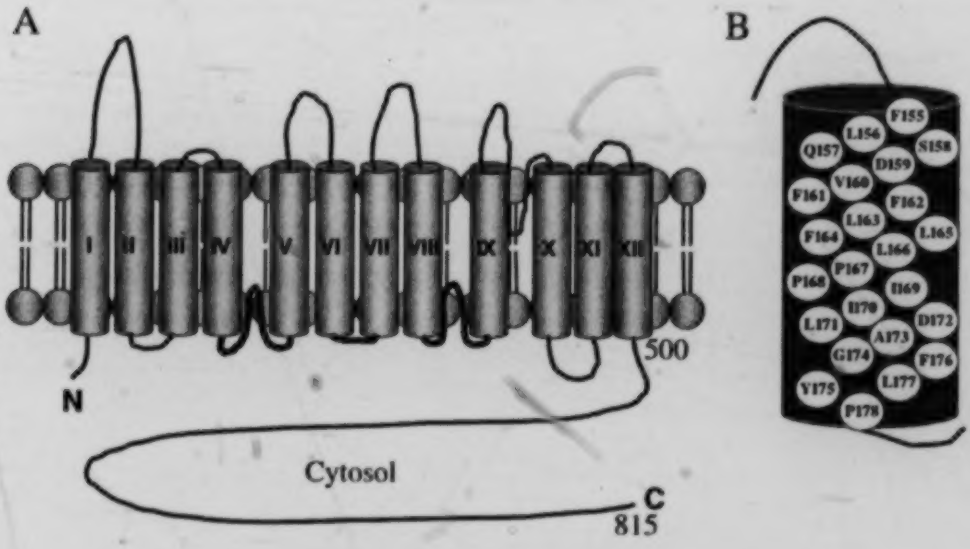

A
B
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
500
N
Cytosol
C
815
F155
L156
Q157
S158
D159
V160
F161
F162
L163
F164
L165
L166
P167
P168
I169
I170
L171
D172
A173
G174
Y175
F176
L177
P178

Figure B-1. **Topological model of the NHE1 isoform of the Na^+/H^+ exchanger.** The orientation of TM segments I-XII of the NHE1 isoform of the Na^+/H^+ exchanger is illustrated in panel (A) (6). Crude model of amino acids present in TM IV is shown in panel (B).

Figure B-1 (above) illustrates a general model of the Na^+/H^+ exchanger (Figure B-1A) and a schematic model illustrating TM IV (Figure B-1B). To examine which amino acids of TM IV are critical to activity of the Na^+/H^+ exchanger and which amino acids are pore-lining, we used the cysteineless Na^+/H^+ exchanger (cNHE). Each residue in TM IV of cNHE1 was individually mutated to a cysteine residue. Initial experiments examined whether these mutant forms of the Na^+/H^+ exchanger had activity. Surprisingly, TM IV was exquisitely sensitive to mutation. Figure B-2 shows the effects of mutation of amino acids of TM IV to cysteine on NHE activity. Substitution of any of the amino acids with cysteine in TM IV resulted in significant reductions of the measurable activity of the protein in all mutants ($P<0.05$). In particular residues F155, L156, S158, D159, F162, F164, P167, P168, D172, Y175 and F176 retained less than 20% of the control activity.

These mutants were not used for further characterization of activity. We have previously demonstrated that P178 is not critical for function of NHE1 (10).

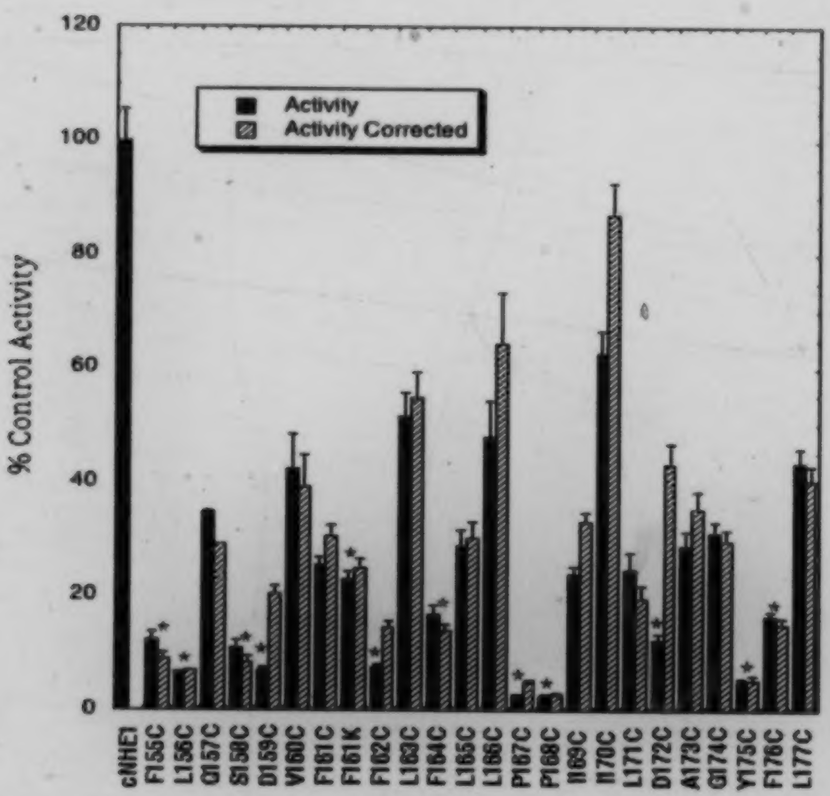


Figure B-2. **Rate of recovery.** An acid load by AP-1 cells stably transfected with cysteineless NHE1 (cNHE1), and Na^+/H^+ exchanger mutants (F155 to L177 individually changed to Cys plus F161K). Na^+/H^+ exchanger activity was measured after transient induction with an acid load as described in the "Materials and Methods". The activity of cNHE1 stably transfected with NHE1 was 3.2 pH units/min, and this value was set to 100%. All mutations to cysteine were done in the background of the cysteineless NHE1 and activities are percent of those of cNHE. The F161K mutation was done in the wild-type NHE1 background and its activity is expressed as a percent of the wild type NHE1 activity (4.25 pH units/min). All results are mean +/- standard error of at least 10 determinations from two independently made stable cell lines. Results are shown for mean activity both uncorrected (black) and normalized for surface processing (hatched). * indicates mutants with unnormalized activity that is less than 20% of cNHE1.

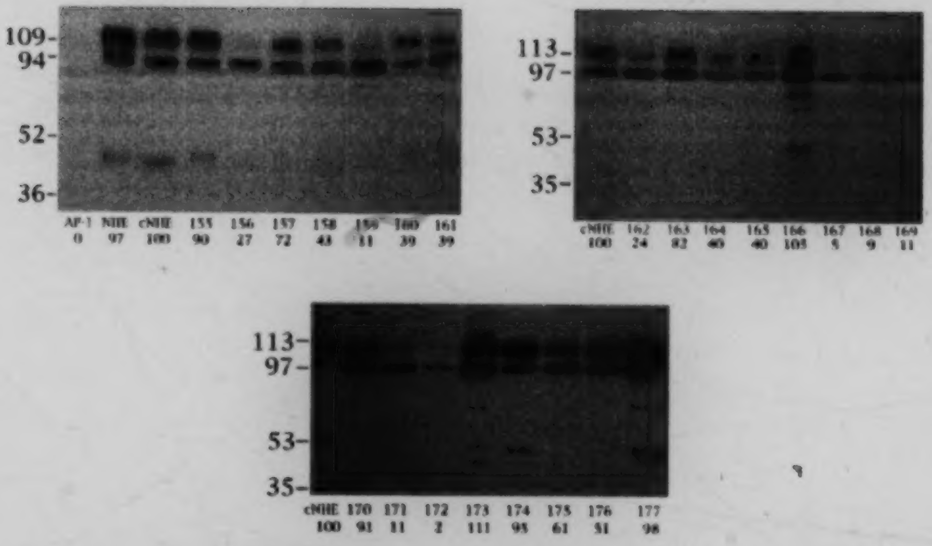

109-
94-
52-
36-
AP-1 NHE cNHE 155 156 157 158 159 160 161
0 97 100 90 27 72 43 11 39 39
113-
97-
53-
35-
cNHE 162 163 164 165 166 167 168 169
100 24 82 40 40 105 5 9 11
113-
97-
53-
35-
cNHE 170 171 172 173 174 175 176 177
100 91 11 2 111 95 61 51 98

Figure B-3. **Western blot analysis of cell extracts from control and stably transfected AP-1 cells.** Cell extracts were prepared from control (AP-1) cells and from cells stably transfected with cDNA coding for HA-tagged: wild type NHE1 (NHE), cysteineless NHE1 (cNHE), F155C, L156C, Q157C, S158C, D159C, V160C, F161C, F162C, L163C, F164C, L165C, L166C, P167C, P168C, I169C, I170C, L171C, D172C, A173C, G174C, Y175C, F176C and L177C. 100 μg total protein was loaded in each lane. Results are typical of at least two experiments. Numbers underneath the lanes indicate the values obtained from densitometric scans of the 110 kDa band relative to cNHE1.

Figure B-3 (above) illustrates a Western blot of total cell extracts from AP1 cells stably expressing the single cysteine mutants. Both the mutant and wild-type exchangers displayed the same pattern of immunoreactive bands, with a larger band at approximately 110-kDa which represents the glycosylated form of the mature Na^+/H^+ exchanger and a smaller band at approximately 95-kDa that represents an immature form of the exchanger that is not fully glycosylated (10). Both the native wild type NHE and the cysteineless NHE1 show strong immunoreactive bands of 110-kDa size and weaker 95-kDa, unglycosylated NHE1. The amount of mature 110-kDa NHE is quantified below each lane, relative to cNHE1. Cells lines containing the mutants F155C, Q157C, L163C, L166C, I170C, A173C, G174C, Y175C, L177C and possibly S158C and F176C have relatively normal levels of expression. Several cell lines expressed mainly or a high proportion of the 95- kDa form of the protein including L156C, D159C, F161C, F162C, P167C, P168C, I169C, L171C, Y175C and F176C. Several cell lines also expressed much less of the mature protein. These mutants included L156C, D159C, V160C, F161C, F162C, F164C, L165C, P167C, P168C, I169C, L171C and D172C.

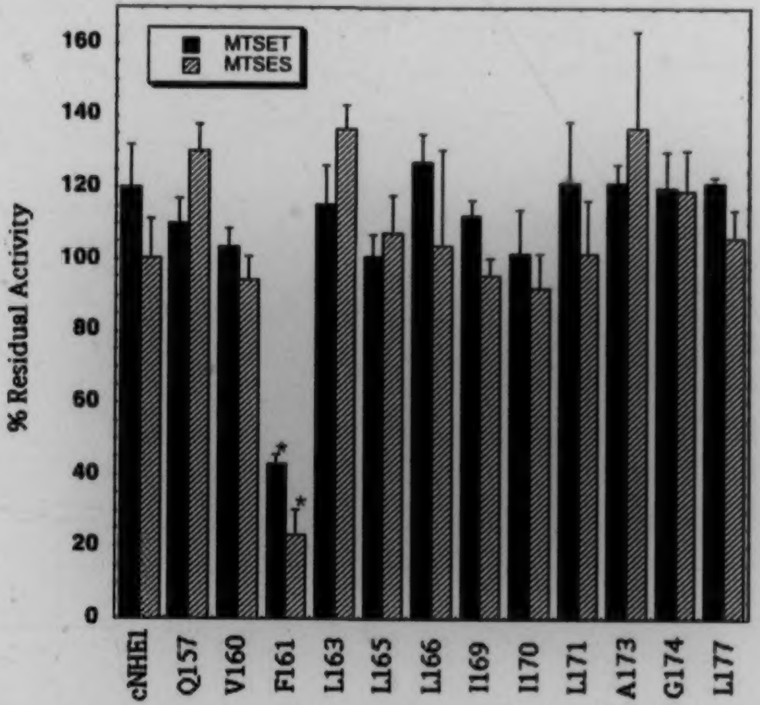


Figure B-4. **Effects of MTSET and MTSES on activity of cNHE1 and single cysteine mutant NHE1 containing cell lines.** Na^+/H^+ exchanger activity was measured after transient induction with an acid load as described in the "Materials and Methods". Cells were subsequently treated with 10 mM reagent before a second transient acidification. Results illustrated represent the % of activity after the second acid load, in comparison to the first. * indicates that the second recovery from acid load was significantly lower than the first at $P < 0.01$, Mann-Whitney U test. Solid filled bars represent MTSET treatments and hatched bars MTSES treatments.

We examined the sensitivity to MTSET or MTSES of the mutant Na^+/H^+ exchangers that had greater than 20% residual activity of the cysteineless NHE1 (Figure B-4). Of the active Na^+/H^+ exchangers, only F161C was affected by treatment with either MTSET or MTSES. This resulted in decreases in Na^+/H^+ exchanger activity of approximately 60% and 80% respectively. To confirm that F161 was a critical residue we created the mutant F161K which introduced a positive charge into this location in the wild-type NHE1 background. This mutant retained only approximately 23% of wild-type NHE activity (Figure B-2), had reduced expression (not shown) and reduced surface processing (*Table B-1*) relative to wild-type NHE1 supporting the importance of F161 in NHE1 function.

Cell line	Plasma Mem (% of total)
wNHE1	76.9 +/- 5.2
cNHE1	50.6 +/- 6.8
F155C	68.9 +/- 5.4
L156C	47.5 +/- 1.4
Q157C	60.6 +/- 2.0
S158C	65.3 +/- 2.2
D159C	17.3 +/- 3.8*
V160C	54.6 +/- 3.6
F161C	42.2 +/- 2.5
F161K	46.6 +/- $6.4^{\dagger,*}$
F162C	26.5 +/- 8.2
L163C	47.5 +/- 4.6
F164C	61.1 +/- 0.8
L165C	48.2 +/- 9.4
L166C	37.5 +/- 3.9
P167C	28.2 +/- 9.9
P168C	47.3 +/- 5.3
I169C	36.6 +/- 7.1
I170C	36.4 +/- 2.0
L171C	64.1 +/- 4.2
D172C	14.5 +/- 1.3*
A173C	41.5 +/- 5.6
G174C	52.8 +/- 10.2
Y175C	46.8 +/- 2.2
F176C	54.4 +/- 3.9
L177C	54.3 +/- 7.3

Table B-1. Summary of subcellular localization in AP-1 cells of wild type (wNHE1), cysteineless NHE1 (cNHE1), and of TM IV Na^+/H^+ exchanger mutants. Analysis of localization was as described in the "Materials and Methods". The percent of the total NHE1 protein that was found within the plasma membrane (Plasma Mem) is indicated. The results are mean +/- SE of at least 3 determinations. * indicates significantly reduced in comparison to cNHE1 at $P < 0.05$. † indicates significantly reduced in comparison to wNHE1 at $P < 0.05$.

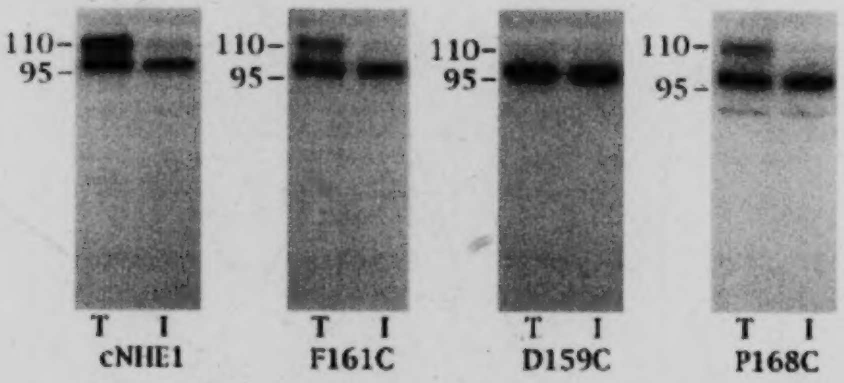


Figure B-5. **Example of a Western blot for determination of localization of the Na⁺/H⁺ exchanger.** Sulfo-NHS-SS-biotin treated cells were lysed and solubilized proteins were treated with streptavidin agarose beads to bind labeled proteins as described in the "Materials and Methods". Equivalent samples of total cell lysate (T) and unbound lysate (I, intracellular) were run on SDS-PAGE. Western blotting with anti-HA antibody identified NHE1 protein. cNHE1, cysteineless NHE1; representative mutants are F161C, D159C and P168C.

We have earlier found that mutation of amino acids of the Na⁺/H⁺ exchanger sometimes causes the protein to be targeted to an intracellular location (*10, 25*). We therefore used quantitative measurement of the intracellular localization of NHE1 within AP-1 cells. Cells were treated with sulfo-NHS-SS-biotin, then lysed and solubilized, and labeled proteins were bound to streptavidin-agarose beads. An equal amount of the total cell lysate and unbound lysate was separated by size using SDS-PAGE followed by Western blotting with anti-HA antibody to identify tagged NHE1 protein. Figure B-5 illustrates examples of the results and Table B-1 summarizes the results quantitatively. In Figure B-5, the total lanes (T) and intracellular lanes (I) illustrate the NHE1 protein present in these samples. The intracellular lane (I) was that fraction of the Na⁺/H⁺ exchanger that did not bind to the streptavidin-agarose beads. Cysteineless NHE1, and most of the mutant exchanger proteins were predominantly present on the plasma membrane in similar amounts. Na⁺/H⁺ exchanger mutants D159C and D172C were significantly reduced in expression on the plasma membrane of the cell. Correcting the activity of these mutants for changes in targeting (Figure B-2) demonstrates that a significant amount of the apparent reduction in activity was due to mistargeting of these proteins. Nevertheless both still retained less than half of the activity of the cysteineless NHE1 protein. Correcting for targeting did not greatly change the values for activity of other mutants with the next largest changes being increases in activity of about 20% for L166C and I170C proteins.

B.3 Production and characterization of TM IV

To examine the structure of TM IV we produced it as a fusion protein with GB-1. On SDS-PAGE the fusion protein appeared as a band of approximately 8-kDa in size (not shown). This was slightly less than the predicted size of 11 kDa. We confirmed the identity of the protein by Mass spectrometry, suggesting that the protein ran with a somewhat anomalous weight, typical of many membrane proteins. The yield of the GB-1-TMIV fusion protein was typically approximately 25 mg/L of cells. After CNBr treatment to free TM IV, it was then purified by reverse phase HPLC. We confirmed the identity of TM IV by Mass spectrometry. For the unlabelled TM4 the expected mass (with a C-terminal homoserine lactone) is 3138.8 Da and the observed mass in retained preparative HPLC fractions ranged over 3138.4-3134.0 Da (parent peak mass differed slightly from fraction to fraction; ionic adducts also observed). Impurities at mass 4405-4411 Da and 9910-9915 Da were detected in less than 25% of the total pooled fractions each, at intensities ranging from 0.5-2% relative to the parent peak plus ionic adducts. For the uniformly 15N labeled peptide, the expected mass was 3165.8, and observed mass range was 3165.6-3171.6 Da. Impurities at mass 3085-3086 Da, 4428-4430 Da, and 4816-4820 Da were detected in less than 10% of the total pooled fractions each, at intensities ranging from 1-4% relative to the parent peak plus ionic adducts. Purity was therefore estimated to be greater than 95% for each of the peptides after pooling of the HPLC fractions.

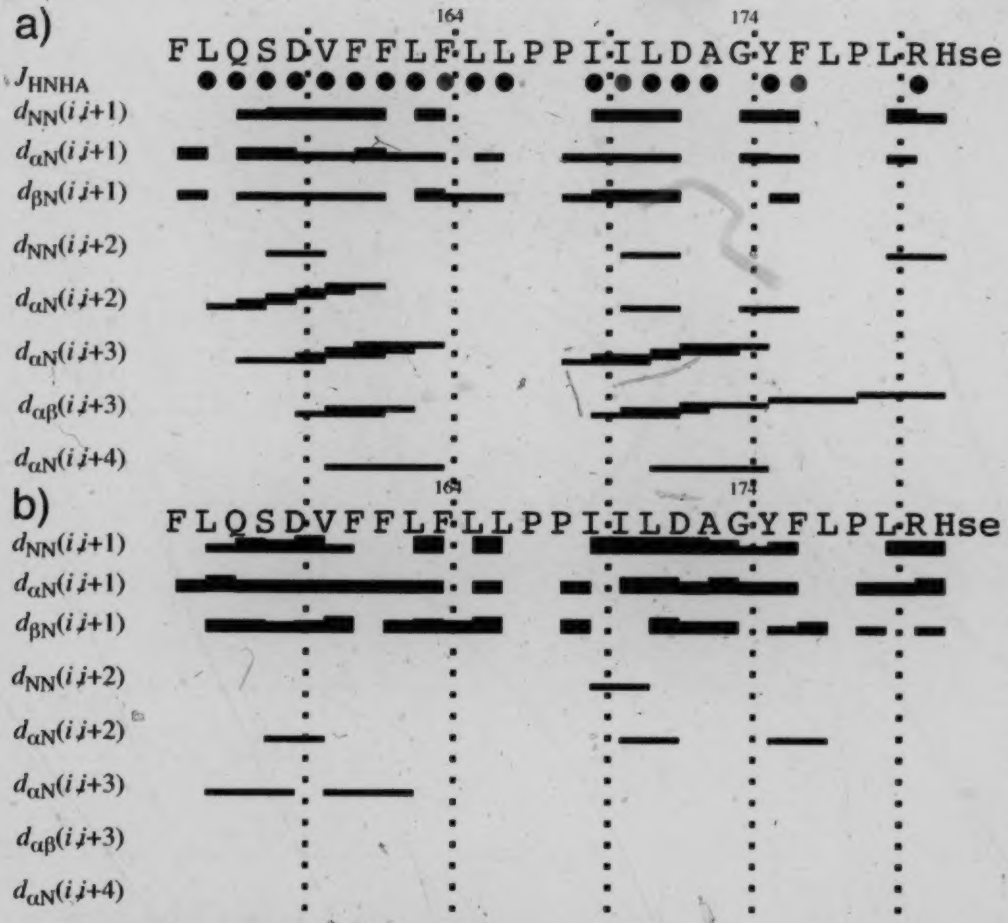


Figure B-6. **Summary of observed NOE's.** NOE connectivities (150 ms mixing time) observed between the indicated pairs of H^N, H^α, and H^β protons of the TM4 segment in [a] $CD_3OH:CDCl_3:H_2O$ (4:4:1 v:v:v) and [b] DMSO. Dashed lines separate groups of 5 residues. Residues for which J_{HNHA} values were usable in $CD_3OH:CDCl_3:H_2O$ are indicated by circles colored black for ± 2 Hz restraints or grey for ± 2.5 Hz restraints. Figure modified from CYANA (L.A. Systems Inc.) output.

B.4 Choice of NMR conditions

To determine the structure of TM IV we tested several solvent mixtures. The TFE:H_2O mixtures and methanol appeared more promising than either the $CD_3OH:CDCl_3:H_2O$ mixture or DMSO, in that wider dispersions of chemical shifts were observed in the backbone H^N (7-9 ppm) and H^α (3.5-5 ppm) regions of 1D ^1H spectra (data not shown). Unfortunately, the peptide precipitated out of each of these solutions within days, making multidimensional NMR data acquisition impractical. In the range of 5-40°C, 30°C gave rise to well resolved, minimally broadened ^1H spectra (data not shown). Therefore, structural studies were carried out in $CD_3OH:CDCl_3:H_2O$ and DMSO at 30°C. These ~2 mM peptide samples with 0.5 mM DSS as the internal chemical shift reference are stable at room temperature, providing reproducible spectra over the course of several months.

B.5 Resonance assignment

Standard sequential assignment methods were applied (26), using TOCSY and DQF-COSY experiments along with H^N values observed by natural abundance ^{15}N-HSQC experiments for spin-system assignment and NOESY connectivity data to walk through the sequence. This provided assignment of all backbone H^N and H^α, except for the N-terminal H^N, all side-chain H^β resonances, and most other side-chain proton resonances. The C- terminal residue was identified as a homoserine lactone (Hsl) rather than free homoserine, since its H^α and side-chain ^1H chemical shifts correspond to α-amino-γ-butyrolactone rather than to L-homoserine (27). This was upheld by the molecular weight indicated by MALDI Mass Spectrometry of the purified peptide when considering the fact that Hsl residues are not undergo subsequently conversion to the free homoserine moiety, since this reaction is reported to be very slow at the peptide C-terminal even with catalytic basic conditions (28). Due to spectral overlap even at 800 MHz, caused primarily by repetition in the primary sequence, a few side-chain protons could not be resolved. In $CD_3OH:CDCl_3:H_2O$, aromatic ring protons could not be identified, H^δ shifts for residues L165, P167, I169, P178, and L179 and H^γ of P167 were unresolved. In DMSO, all aliphatic side-chain resonances were resolved, and all ring H^δ resonances; in the case of F155 and Y175, H^ϵ resonances were also resolved. Note that frequent overlap of H^β and H^γ resonances was observed with Leu and Ile side-chains. ^{15}N-edited TOCSY and NOESY experiments with uniformly labeled peptide samples confirmed spin system and resonance assignment, and allowed several additional protons to be assigned. Finally, natural abundance ^{13}C-HSQC experiments allowed identification of all C$^{\alpha}$ and C$^{\beta}$ resonances except C$^{\alpha}$ of the C-terminal Hsl in $CD_3OH:CDCl_3:H_2O$ and C$^{\beta}$ of P167 in DMSO All resonance assignments have been deposited in the BMRB.

B.6 Structure calculation

Homonuclear NOESY experiments were used exclusively to define NOE contacts between protons. A variety of mixing times were examined under each solvent condition (100-450 ms; data not shown). In each of the solvents, the 150 ms mixing time provided peaks of sufficient intensity to analyze without apparently extensive spin diffusion. Therefore, this mixing time was employed for quantitative peak volume analysis. A large number of intra- and inter-residue contacts were observed in each solvent condition. In DMSO, 302 intra-molecular, 244 sequential and 62 medium range contacts were unambiguously assigned. In $CD_3OH:CDCl_3:H_2O$, there were 222 intra-residue, 140 sequential, and 96 medium range unambiguous contacts assigned. More contacts were able to be assigned in the DMSO conditions due to the additional Phe and Tyr ring proton assignments. In $CD_3OH:CDCl_3:H_2O$, the TM IV sample showed two regions with significant i,i+3 contacts, often indicative of α-helical character, which were not apparent in DMSO (Figure B-6). The first of these regions is coincident with a large number of i,i+2 contacts, however, which together indicate potential β-turns. Chemical shifts have not at present been incorporated into the structural calculations because a quantitative comparison to random coil or BMRB average chemical shifts observed in aqueous solution may be altogether incorrect. ^{3}J-coupling constants extracted from ^{15}N-edited HNHA experiments were used where possible to aid in specifying ϕ-dihedral angles during the simulated annealing protocol. Since the peptide was less structured in DMSO than in $CD_3OH:CDCl_3:H_2O$, as evinced by the greatly reduced number of medium range NOE contacts (Figure B-6), we analyze only the $CD_3OH:CDCl_3:H_2O$ structural features in detail at present. Characteristics of the final ensemble of retained structures are given in *Table B-2* below

Non-redundant NOE restraints	
Total	458
Intra-residue	222
Sequential	140
Medium range ($\lvert i-j \rvert \le 4$)	96
Energies (kcal/mol)	
E_{Total}	106.8±9.8
E_{NOE}	8.67±2.04
Restraint violations	
$0.3\ \text{Å} > \lvert \text{NOE} - \text{distance} \rvert > 0.2\ \text{Å}$	12
$\lvert \text{NOE} - \text{distance} \rvert \ge 0.3\ \text{Å}$	0
J_{HNHA} violations	81

Table B-2. Structural statistics for the final ensemble of 600 structures out of the 1000 calculated for $CD_3OH:CDCl_3:H_2O$.

DISCUSSION

B.1 Functional analysis of TM IV

Transmembrane segment four (TM IV; residues 155-177) of the NHE1 isoform of the Na^+/H^+ exchanger has been implicated in the ion transport and inhibitor binding properties of the protein (7-9). We have recently shown (10) that the double proline pair P167 and P168 is critical to the function of NHE1. They were suggested to be critical in maintenance of an appropriate structure of TM IV and necessary for normal NHE1 transport function. One method to determine the functional role of individual amino acids of a transmembrane segment of a membrane protein is cysteine-scanning mutagenesis in combination with reaction with sulfhydryl reagents. Cysteine-scanning mutagenesis takes advantage of the fact that the sulfhydryl moiety is the most reactive functional group in a protein (29). It has been used to determine pore-lining residues in numerous membrane proteins (30-32) including the lactose permease of *E. coli* (33), the mouse acetylcholine receptor (34), the human glucose transporter Glut1, and the human anion exchanger isoform 1 (AE1) (35). Cysteine-scanning mutagenesis uses the highly reactive sulfhydryl moiety to determine accessibility of side chains of amino acids. Two sulfhydryl reactive reagents that are (36) often used in these studies are MTSET and MTSES which are membrane impermeant (*37, 38*), react with pore lining residues surrounded by water, and cannot reach with residues within the hydrophobic bilayer.

Initial experiments substituted each of the amino acids in TMIV of cysteineless NHE1 with cysteine residues. We found that TM IV was exquisitely sensitive to mutation. Twelve of the cysteine mutants had less than 20% of the activity of the control cNHE1. Western blotting revealed that in some cases the reduction in activity was due to a lower level of expression of the protein (Figure B-3), however in most cases correction for the amount of plasma membrane protein (*Table B-1*, Figure B-2) did not explain the loss of activity. Several of the mutants were expressed mainly as unglycosylated protein and many, but not all of these had greatly reduced activity. We have previously found that mutation of some amino acids can greatly affect the glycosylation levels of the proteins. We examined if these mutations affected the surface targeting of the protein. Our findings, that many of the mutations of TM IV affected targeting, expression or activity, are similar to the results found with the human anion exchanger (35) and with TM XI of lactose permease (33). In contrast, for the tetracycline resistant transporter of *E. coli* (39) and in P-glycoprotein (40), it was possible to mutate all the amino acids of a TM segment and always retain appreciable activity. This susceptibility to mutation appears to vary not only between proteins, but also within different transmembrane segments of the same protein. TM XI of lactose permease had several sensitive amino acids (33), while TM XII only had one (41). This property could be reflective of both the importance of the residues in the particular segment, and of the importance of the segment itself to the structure and function of the protein. In our case, it was clear that there was a relatively strict requirement for many of the amino acids of TM IV, and this could be indicative of precise structural and functional roles.

Treatment with both positively charged MTSET and negatively charged MTSES revealed that Phe161 is accessible to and reacts with these sulfhydryl-reactive reagents (Figure B-4). The most likely explanation for this reactivity is an interaction at a site that lines the ion translocation pore. The reagents then at least partially block the pore and inhibit ion transport (35). Thus, Phe161 in TM IV is the first residue to be unambiguously identified as lining the ion transport pore of NHE1. Mutation of Phe161 to Lys resulted in loss of most of NHE1 activity (Figure B-2) confirming that this is a critical amino acid.

B.2 Structural analysis of TM IV

We examined the structure of TM IV in $CD_3OH:CDCl_3:H_2O$ and in DMSO at 30°C. As with many peptides studied by NMR (*20, 42*) the TM IV segment does not, as a whole, assume a single conformation. Rather, sections within the peptide converge structurally. Consideration of these sections relative to each other then provides insight into the TM IV segment overall. In a given state of the intact NHE1 protein, a single conformation would presumably be preferred by the TM IV segment. However, the structures that we present here represent the most energetically favorable forms of this TM segment in a low dielectric environment mimicking both a lipid bilayer and a protein interior. Therefore, it is reasonable to presume that they would also be favorable within the setting of the NHE1 protein. Indeed, a number of studies have demonstrated strong correspondence between structures of peptides or protein segments obtained in membrane mimetic solvents to structures obtained by solution state NMR in micelles or to entire membrane proteins determined by X-ray crystallography (11-13). Interestingly, we find that $CD_3OH:CDCl_3:H_2O$ provides a well defined peptide structure, while the peptide in DMSO appears much less structured. This implies that there is a great deal of value in determining membrane peptide structures in multiple solvent conditions, since a single solvent may not induce sufficient sampling of the structured state.

Three sections of four to nine residues of the TM IV segment converge structurally and were defined as core regions by NMRCORE (23) and confirmed by analysis of RMSD values (*Table B*-3). Namely, the stretches of residues D159-L163, L165-P168, and I169-F176 each converge. These are illustrated in Figure B-7(a-c), and their relationship is demonstrated schematically in Figure B-7d. Each section superposes extremely well, with all convergent residues identified in Figure B-7d for the 600-member ensemble having a C^{α} RMSD relative to the average structure in the range of 0.25-0.85 Å. The heavy backbone and side-chain atom RMSD's are < 1.5 Å over D159-L163 and I169-G174 as well as for L165 and P167 (*Table B-3*), which implies good structural convergence at the side-chain level for these segments. The boundary points of converged stretches are also very clear upon examination the RMSD values in Table 3: S158 and Q157 rapidly diverge from D159-L163; F164 has a much higher RMSD than the surrounding residues when the peptide structures are superposed at either 159-163 or 165-168; P168 clearly fits well with 165-168, but not with 169-176, while the converse is true of I169; finally, L177, P178 and L179 demonstrate large jumps up in RMSD values relative to the superposition 170-178. One of the ensemble members is shown in Figure B-7e as a representative structure. Given that the segment must span a membrane with width on the order of 25-35 Å, a relatively extended ensemble member such as that pictured in Figure B-7e is more representative of the expected biological configuration than one that curls back upon itself. This particular structure must be considered as only demonstrative of the relation between the convergent segments of the TM IV segment, rather than as a representative of any global structuring.

The first feature of note is that the TM IV segment certainly does not resemble a canonical transmembrane α-helix. This is despite the fact that there are helix-capping sequences inherent within TM IV over the segment studied. The N-terminal sequence FLQSDV (155-160) fits well with a type Ib cap, while the sequence LDAGYF (170-176) fits a type Va C-terminal cap, with capping nomenclature as defined by Aurora and Rose (*45*). These cap sequences would imply that a helix should form between L156 and A173. However, the NMR data readily demonstrate a lack of helicity, even at relatively long NOESY mixing times. Contiguous (i, i+3) and (i, i+4) contacts indicative of an α-helix (26) are not observed. In fact, as implied in Figure B-7d, only a very short stretch of the segment even approaches a helical structure. This should not be taken imply to a lack of structure: in fact, eight type IV β-turns occur in a major proportion of ensemble members, with four found in more than 90% of the structures. Note that this means a frequent observance of "disallowed" ϕ/ψ angles would actually be expected (*46*), and, correspondingly, a PROCHECK-NMR (*47*) analysis places 12.2% of the non-Pro and Gly residues in the disallowed region of the Ramachandran plot.

Residue	RMSD - C^{α}	RMSD – all heavy atoms
Q157[a,d]	1.83	3.43
S158[a,d]	1.23	1.86
D159[a]	0.43	1.24
V160[a]	0.25	0.59
F161[a]	0.30	1.42
F162[a,d]	0.45	1.50
L163[a]	0.53	1.16
F164	1.54[a] / 2.49[b]	2.65[a] / 4.00[b]
L165[b]	0.80	1.40
L166[b]	0.57	1.94
P167[b]	0.38	0.85
P168	0.85[b] / 1.43[c]	1.55[b] / 1.81[c]
I169[d]	3.16[b] / 0.76[c]	4.36[b] / 1.34[c]
I170[c]	0.70	1.24
L171[c]	0.59	1.41
D172[c]	0.48	1.32
A173[c]	0.52	0.64
G174[c,d]	0.55	0.56
Y175[c]	0.36	2.00
F176[c]	0.70	1.77
L177[c,d]	1.24	2.61
P178[c]	1.36	1.64
L179[c,d]	2.66	3.53

Table B-3. RMSD for C^{α} and for all heavy atoms of the 600 lowest energy structures of the 1000 member ensemble of the TM IV peptide superposed over each convergent stretch and given relative to the average structure for that superposition. Residues outside converged segments and at borders between convergent segments are also shown.

a) Superposed on C^{α} of 159-161 and 163 using LSQKAB(*43, 44*).
b) Superposed on C^{α} of 165-168.
c) Superposed on C^{α} of 170-173, 175-176, and 178.
d) Not defined as core using program NMRCORE.

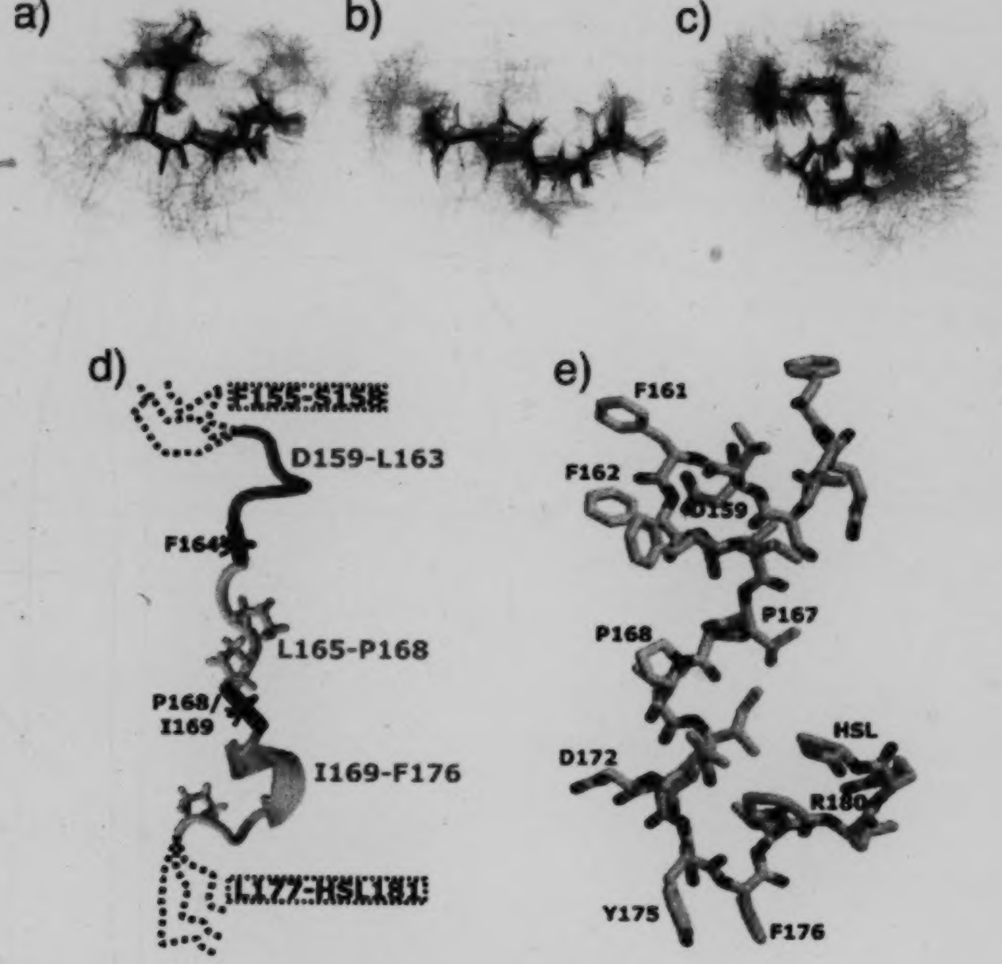

a)
b)
c)
d)
F155-S158
D159-L163
F164
L165-P168
P168/
I169
I169-F176
L177-HSL181
e)
F161
F162
D159
P167
P168
HSL
D172
R180
Y175
F176

Figure B-7. **The NMR Structure of TM IV.** Convergent structural motifs pinpointed by NMRCORE (23) in the lowest energy 600 members of the final 1000 structure ensemble satisfying NOE and J-coupling restraints determined for the TM4 peptide in CD_3OH:$CDCl_3$:H_2O (4:4:1 v:v:v). [a]-[c] Show the lowest energy 60 structures, with backbone atoms colored black and side-chains grey. Convergent [a] loop structure over residues 159-163; [b] defined extended structure over 165-168; and, [c] helical stretch at 169-176. [d] Schematic demonstrating convergent stretches (grey) in relation to pivot residues F164 and P168/I169 (black) and flexible N- and C-termini (dashed lines). Note that only proline side-chains are drawn. [e] Stick diagram with CPK coloring of a single member of the ensemble (*i.e.* the flexible linkers and termini in [d] have assumed a discrete configuration) demonstrating the overall extended nature of the peptide.

Of the three core regions defined by convergent backbone atoms, residues D159-L163 are the most affected by these β-turns. D159 is the i+2 residue in one type IV turn and the first residue in a second, V160 is the i+3 and i+1 residue for these turns, as well as the first residue in a third turn; following along the sequence, F161 and F162 both participate in the turns initiated at D159 and V160, and L163 participates as the i+3 residue for V160. Therefore, this series of turns, observed in 81-94% of the ensemble, serves to converge this five residue stretch. Interestingly, L165-P168 are nowhere near as strongly influenced by β-turns. Rather, these residues appear to be quite extended (Figure 7b), with P168 initiating type I or IV β-turns in ~ 85% of the ensemble. Finally, residues I169-F176 converge strongly, providing the only appreciable α-helical character over the entire transmembrane segment. The helical stretch extends over 4-6 residues for ~77% of the ensemble members, starting between I169 and L171 and ending at G174, with each one of these members of the ensemble showing helical character over the 4 residue segment of L170-G174. The C-terminal of this convergent stretch is defined by G174 and Y175, which each act as the first residue for a sequential pair of type IV β-turns in 92% or more of the ensemble members.

While similar experimental resolution in NMR studies has been achieved using micelles in aqueous solution(*48*), such an environment would surround the TM IV peptide with hydrophobic detergent tails. In its natural setting, this segment of the NHE1 protein would contact lipid tail groups, surrounding polypeptide segments, and a pore region with relatively high dielectric constant. In the case where a transmembrane segment has a very clearly defined structure, such as a canonical membrane spanning α-helix, the micellar environment serves as an excellent conformational stabilizer. In the micellar state, we would anticipate that the allowed rotations about F164 and P168/I169 would be greatly reduced. However, this may actually be an artificial constraint upon the peptide structure relative to the solvent setting, which allows increased conformational sampling. We feel that the present structure in a mixture of solvents representing a variety of dielectrics provides a reliable representation of the best-structured regions of the TM IV segment. This should be useful in interpretation of future experiments using more complete sections of NHE1, in particular for solid state NMR experiments which are aided by a knowledge of the most likely polypeptide backbone configuration (*49*).

B.2 Correlating structural and functional data

In light of the structural data, plausible explanations for the loss in activity observed for some of the most severe cysteine scanning mutants can be advanced. Mutations to P167C and P168C practically abrogate all exchanger activity and mutations of these prolines to other residues did not allow for a return to normally functioning Na^+/H^+ exchanger (10). Mutation from an imino to a much freer amino acid moiety would readily modify the structural motif observed at this pair of prolines (Figure B-7b). Given the loss of activity, it is likely that the extended structural nature of this motif is crucial in NHE1 folding into the functional form. The Y175C mutation also causes almost complete loss in activity. Interestingly, Y175 is observed to participate as the second residue in an extremely well converged double (*i*, *i*+1) type IV β-turn, just N-terminal to the aromatic residue F176. Note that F176C also causes significant loss of activity. Analysis of the torsion angles between defined between the C^α-C^β, -C^γ, and -C^ξ vectors of Y175 and F176 for the ensemble of 600 structures show that the aromatic side-chain moieties of Y175 and F176 have a strong tendency to fall within ~50° of each other relative to the polypeptide backbone over this region of the TM IV segment. Since these are large, hydrophobic side-chains, it seems likely that they would be at the interface with neighboring transmembrane segments or protruding into the lipid bilayer rather than directly into the channel environment. Substitution of these aromatic groups with the significantly less bulky, polar cysteine side-chain would disrupt these interactions, thereby disrupting the NHE1 structure as a whole, leading to the observed loss in activity. However, the mutated F161C side-chain does appear to be pore-exposed, so this is not a certainty.

The mutation F161C allows for reactivity with the sulfhydryl-reactive reagent MTSET that inhibits Na^+/H^+ exchanger activity. This implies that residue 161 is facing the NHE1 translocation pore. With this in mind, examination of the structure of the TM IV segment provides a very interesting revelation. F161 side-chain is actually fully exposed when put in the context of residues F155-I169 in every member of the ensemble (data not shown). Since rotation of the structurally convergent sections is observed about F164 (Figure B-7d), any further analysis of the two C-terminal domains relative to F161 would be purely speculative and it was therefore not possible to determine at this stage which other C-terminal amino acids face the translocation pore. However, the exposed nature of the F161 side-chain at the extracellular end of the segment is demonstrated in all 600 retained ensemble members satisfying the experimental constraints, meaning that the isolated TM IV segment in membrane mimetic solvent assumes a structure consistent with the reactivity observed for the exchanger as a whole. Over the convergent stretch of D159-L163 (Figure B-7a), the D159 side-chain is uniformly observed on the same face of the peptide. Therefore, it seems likely that the anionic D159 side-chain would also protrude into the pore of the exchanger, allowing it to act in cation coordination during translocation. Mutation of this amino acid eliminated activity and reduced expression of the protein, consistent with a critical role for this amino acid. We have previously postulated that polar residues within transmembrane segments of the Na^+/H^+ exchanger are important in cation coordination and transport (*50, 51*), and the structure of TM IV supports such a role for this polar residue.

CONCLUSIONS

Our study has given the first detailed structural and functional information on TM IV of the NHE1 isoform of the Na^+/H^+ exchanger. TM IV appears to be very sensitive to alterations in its amino acid sequence. We show that while it is a well-structured transmembrane segment, its structure is uniquely different from a typical α-helix. Possibly this reflects its pivotal role in cation binding and transport. F161 and possibly D159 are amino acid residues that protrude into the pore of the channel.

REFERENCES

1. Orlowski, J., and Grinstein, S. (1997) *J. Biol. Chem. 272*, 22373-22376.
2. Grinstein, S., Rotin, D., and Mason, M. J. (1989) *Biochim. Biophys. Acta 988*, 73-97.
3. Pouyssegur, J., Sardet, C., Franchi, A., L'Allemain, G., and Paris, S. (1984) *Proc. Natl. Acad. Sci. USA 81*, 4833-4837.
4. Fliegel, L. (2001) *Basic Res. Cardiol. 96*, 301-305.
5. Mentzer, R. M., Jr., Lasley, R. D., Jessel, A., and Karmazyn, M. (2003) *Ann Thorac Surg 75*, S700-S708.
6. Wakabayashi, S., Pang, T., Su, X., and Shigekawa, M. (2000) *J. Biol. Chem. 275*, 7942-7949.
7. Counillon, L., Franchi, A., and Pouyssegur, J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4508-4512.
8. Counillon, L., Noel, J., Reithmeier, R. A. F., and Pouyssegur, J. (1997) *Biochemistry 36*, 2951-2959.
9. Touret, N., Poujeol, P., and Counillon, L. (2001) *Biochemistry 40*, 5095-5101.
10. Slepkov, E. R., Chow, S., Lemieux, M. J., and Fliegel, L. (2004) *Biochem. J. 379*, 31-38.
11. Girvin, M. E., Rastogi, V. K., Abildgaard, F., Markley, J. L., and Fillingame, R. H. (1998) *Biochemistry 37*, 8817-8824.
12. Katragadda, M., Alderfer, J. L., and Yeagle, P. L. (2001) *Biophys J 81*, 1029-36.
13. Yeagle, P. L., Choi, G., and Albert, A. D. (2001) *Biochemistry 40*, 11932-7.
14. Lindhout, D. A., Thiessen, A., Schieve, D., and Sykes, B. D. (2003) *Protein Sci 12*, 1786-1791.
15. Huth, J. R., Bewley, C. A., Jackson, B. M., Hinnebusch, A. G., Clore, G. M., and Gronenborn, A. M. (1997) *Protein Sci 6*, 2359-64.
16. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) *J Biomol NMR 6*, 135-140.
17. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J Biomol NMR 6*, 277-293.
18. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallographica Section D-Biological Crystallography 54*, 905-921.
19. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772-7777.
20. Wang, J. J., Hodges, R. S., and Sykes, B. D. (1995) *J Am Chem Soc 117*, 8627-8634.
21. Garrett, D. S., Kuszewski, J., Hancock, T. J., Lodi, P. J., Vuister, G. W., Gronenborn, A. M., and Clore, G. M. (1994) *Journal of Magnetic Resonance Series B 104*, 99-103.
22. Kleywegt, G. J. (1995) *CCP4/ESF-EACBM Newsletter on Protein Crystallography 31 (June 1995)*, 45-50.
23. Kelley, L. A., Gardner, S. P., and Sutcliffe, M. J. (1997) *Protein Eng 10*, 737-741.
24. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci 5*, 212-220.

25. Murtazina, B., Booth, B. J., Bullis, B. L., Singh, D. N., and Fliegel, L. (2001) *Eur. J. Biochem. 268*, 1-13.
26. Wuthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley, New York.
27. Pouchert, C. J. (1983) *The Aldrich library of NMR spectra.*, 2nd ed., Aldrich Chemical Company, Milwaukee, WI.
28. Ambler, R. P. (1965) *Biochem J 96*, 32P.
29. Mattson, G., Conklin, E., Desai, S., Nielander, G., Savage, M. D., and Morgensen, S. (1993) *Mol. Biol. Rep. 17*, 167-183.
30. Yan, R. T., and Maloney, P. C. (1995) *Proc Natl Acad Sci U S A 92*, 5973-5976.
31. Slotboom, D. J., Konings, W. N., and Lolkema, J. S. (2001) *J Biol Chem 276*, 10775-10481.
32. Doering, A. E., Nicoll, D. A., Lu, Y., Lu, L., Weiss, J. N., and Philipson, K. D. (1998) *J Biol Chem 273*, 778-83.
33. Dunten, R. L., Sahin-Toth, M., and Kaback, H. R. (1993) *Biochemistry 32*, 12644-12650.
34. Akabas, M. H., Stauffer, D. A., Xu, M., and Karlin, A. (1992) *Science 258*, 307-310.
35. Tang, X.-B., Kovacs, M., Sterling, D., and Casey, J. R. (1999) *J. Biol. Chem. 274*, 3557-3564.
36. Stauffer, D. A., and Karlin, A. (1994) *Biochemistry 33*, 6840-6849.
37. Holmgren, M., Liu, Y., Xu, Y., and Yellen, G. (1996) *Neuropharmacology 35*, 797-804.
38. Liu, J., and Siegelbaum, S. A. (2000) *Neuron 28*, 899-909.
39. Kimura-Someya, T., Iwaki, S., and Yamaguchi, A. (1998) *J. Biol. Chem. 273*, 32806-32811.
40. Loo, T. W., and Clarke, D. M. (1999) *J. Biol. Chem. 274*, 35388-35392.
41. He, M. M., Sun, J., and Kaback, H. R. (1996) *Biochemistry 35*, 12909-12914.
42. Spadaccini, R., and Temussi, P. A. (2001) *Cellular and Molecular Life Sciences 58*, 1572-1582.
43. Kabsch, W. (1976) *Acta Crystallographica Section A 32*, 922-923.
44. Collaborative Computational Project, N. (1994) *Acta Crystallographica Section D-Biological Crystallography 50*, 760-763.
45. Aurora, R., and Rose, G. D. (1998) *Protein Sci 7*, 21-38.
46. Hutchinson, E. G., and Thornton, J. M. (1994) *Protein Sci 3*, 2207-2216.
47. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J Biomol NMR 8*, 477-486.
48. Henry, G. D., and Sykes, B. D. (1994) *Methods Enzymol 239*, 515-35.
49. Marassi, F. M. (2002) *Concepts in Magnetic Resonance 14*, 212-224.
50. Dibrov, P., Young, P. G., and Fliegel, L. (1998) *Biochemistry 36*, 8282-8288.
51. Dibrov, P., and Fliegel, L. (1998) *FEBS Lett. 424*, 1-5.

APPENDIX C

Structural and functional characterization of transmembrane segment IV of the NHE1 isoform of the Na$^+$/H$^+$ exchanger - chemical shift and procheck structural analysis

This appendix contains additional information regarding structural characterization of the presented NMR solution structure of transmembrane spanning helix IV (TMIV) of hNHE1, described in Appendix B of this thesis. This work was recently published in May 2005 in the *Journal of Biological Chemistry* [*J Biol Chem.* 2005 May 6 **280(18)**:17863-72]. All chemical shift data are reported as standard output (*.str) from the BioMagResBank (BMRB accession code 6446) and do not reflect the precision of the chemical shifts. All chemical shift data and structural analysis is from the TMIV peptide in a $CD_3OH:CDCl_3:H_2O$ mixture.

Appendix C-1. Assigned ^{1}H and ^{15}N chemical shifts of TMIV of hNHE1

Note: For simplicity during assignments, the residue numbering for TMIV began at residue ***Phe1***, which corresponded to ***Phe155*** within the hNHE1 protein.

```
 1   1   PHE  CA    C   49.586  0.5   1
 2   1   PHE  CB    C   42.513  0.5   1
 3   1   PHE  HA    H    3.951  0.05  1
 4   1   PHE  HB2   H    3.529  0.05  1
 5   2   LEU  CA    C   57.161  0.5   1
 6   2   LEU  CB    C   41.715  0.5   1
 7   2   LEU  HA    H    4.363  0.05  1
 8   2   LEU  HB2   H    1.798  0.05  1
 9   2   LEU  HB3   H    1.799  0.05  1
10   2   LEU  HD1   H    0.903  0.05  1
11   2   LEU  HD2   H    0.957  0.05  1
12   2   LEU  HG    H    1.578  0.05  1
13   2   LEU  H     H    7.479  0.05  1
14   2   LEU  N     N  116.867  0.5   1
15   3   GLN  CA    C   57.582  0.5   1
16   3   GLN  CB    C   28.463  0.5   1
17   3   GLN  HA    H    4.182  0.05  1
18   3   GLN  HB2   H    2.059  0.05  1
19   3   GLN  HB3   H    2.146  0.05  1
20   3   GLN  HE21  H    7.466  0.05  1
21   3   GLN  HE22  H    6.646  0.05  1
22   3   GLN  HG2   H    2.399  0.05  1
23   3   GLN  HG3   H    2.404  0.05  1
24   3   GLN  H     H    8.783  0.05  1
25   3   GLN  N     N  119.307  0.5   1
26   3   GLN  NE2   N  109.086  0.5   1
27   4   SER  CA    C   59.633  0.5   1
28   4   SER  CB    C   63.031  0.5   1
29   4   SER  HA    H    4.289  0.05  1
30   4   SER  HB2   H    3.865  0.05  1
31   4   SER  HB3   H    3.992  0.05  1
32   4   SER  H     H    7.904  0.05  1
33   4   SER  N     N  113.151  0.5   1
34   5   ASP  CA    C   54.502  0.5   1
35   5   ASP  CB    C   37.190  0.5   1
36   5   ASP  HA    H    4.639  0.05  1
37   5   ASP  HB2   H    2.944  0.05  1
38   5   ASP  HB3   H    2.998  0.05  1
39   5   ASP  H     H    8.144  0.05  1
40   5   ASP  N     N  119.096  0.5   1
41   6   VAL  CA    C   64.943  0.5   1
42   6   VAL  CB    C   31.410  0.5   1
43   6   VAL  HA    H    3.802  0.05  1
44   6   VAL  HB    H    2.117  0.05  1
45   6   VAL  HG1   H    0.886  0.05  1
46   6   VAL  HG2   H    0.972  0.05  1
47   6   VAL  H     H    7.785  0.05  1
48   6   VAL  N     N  117.748  0.5   1
49   7   PHE  CA    C   60.306  0.5   1
50   7   PHE  CB    C   38.475  0.5   1
51   7   PHE  HA    H    4.296  0.05  1
52   7   PHE  HB2   H    3.056  0.05  1
53   7   PHE  HB3   H    3.108  0.05  1
54   7   PHE  H     H    7.579  0.05  1
55   7   PHE  N     N  117.317  0.5   1
56   8   PHE  CA    C   55.172  0.5   1
57   8   PHE  CB    C   38.707  0.5   1
58   8   PHE  HA    H    4.285  0.05  1
59   8   PHE  HB2   H    3.157  0.05  1
60   8   PHE  HB3   H    3.235  0.05  1
61   8   PHE  H     H    7.715  0.05  1
62   8   PHE  N     N  115.671  0.5   1
63   9   LEU  CA    C   57.149  0.5   1
64   9   LEU  CB    C   41.504  0.5   1
```

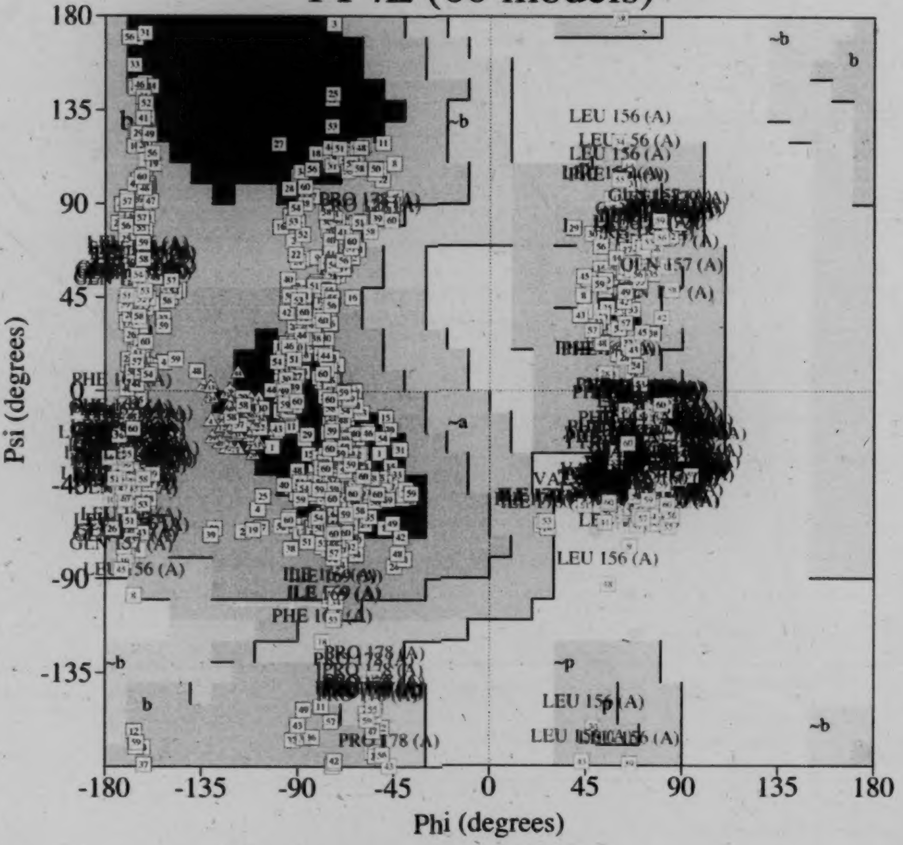
65	9	LEU	HA	H	4.024	0.05	1
66	9	LEU	HB2	H	1.756	0.05	1
67	9	LEU	HB3	H	1.733	0.05	1
68	9	LEU	HD1	H	0.911	0.05	1
69	9	LEU	HD2	H	0.858	0.05	1
70	9	LEU	HG	H	1.410	0.05	1
71	9	LEU	H	H	7.702	0.05	1
72	9	LEU	N	N	116.853	0.5	1
73	10	PHE	CA	C	58.964	0.5	1
74	10	PHE	CB	C	39.023	0.5	1
75	10	PHE	HA	H	4.459	0.05	1
76	10	PHE	HB2	H	3.118	0.05	1
77	10	PHE	HB3	H	3.243	0.05	1
78	10	PHE	H	H	7.524	0.05	1
79	10	PHE	N	N	113.665	0.5	1
80	11	LEU	CA	C	54.704	0.5	1
81	11	LEU	CB	C	40.513	0.5	1
82	11	LEU	HA	H	4.206	0.05	1
83	11	LEU	HB2	H	1.584	0.05	1
84	11	LEU	HB3	H	1.577	0.05	1
85	11	LEU	HG	H	1.306	0.05	1
86	11	LEU	H	H	8.189	0.05	1
87	11	LEU	N	N	120.891	0.5	1
88	12	LEU	CA	C	56.155	0.5	1
89	12	LEU	CB	C	40.918	0.5	1
90	12	LEU	HA	H	4.441	0.05	1
91	12	LEU	HB2	H	1.872	0.05	1
92	12	LEU	HB3	H	1.869	0.05	1
93	12	LEU	HG	H	1.587	0.05	1
94	12	LEU	HD1	H	0.980	0.05	1
95	12	LEU	H	H	7.767	0.05	1
96	12	LEU	N	N	120.365	0.5	1
97	13	PRO	CA	C	64.222	0.5	1
98	13	PRO	CB	C	30.842	0.5	1
99	13	PRO	HA	H	4.409	0.05	1
100	13	PRO	HB2	H	2.252	0.05	1
101	13	PRO	HB3	H	1.895	0.05	1
102	14	PRO	CA	C	65.649	0.5	1
103	14	PRO	CB	C	28.549	0.5	1
104	14	PRO	HA	H	4.274	0.05	1
105	14	PRO	HB2	H	2.285	0.05	1
106	14	PRO	HB3	H	1.956	0.05	1
107	14	PRO	HD2	H	3.559	0.05	1
108	14	PRO	HG2	H	2.021	0.05	1
109	15	ILE	CA	C	64.422	0.5	1
110	15	ILE	CB	C	37.549	0.5	1
111	15	ILE	HA	H	3.829	0.05	1
112	15	ILE	HB	H	2.116	0.05	1
113	15	ILE	HG2	H	0.955	0.05	1
114	15	ILE	HG12	H	1.229	0.05	1
115	15	ILE	HG13	H	1.230	0.05	1
116	15	ILE	H	H	7.363	0.05	1
117	15	ILE	N	N	116.522	0.5	1
118	16	ILE	CA	C	64.067	0.5	1
119	16	ILE	CB	C	37.350	0.5	1
120	16	ILE	HA	H	3.750	0.05	1
121	16	ILE	HB	H	1.972	0.05	1
122	16	ILE	HD1	H	0.905	0.05	1
123	16	ILE	HG12	H	1.334	0.05	1
124	16	ILE	HG13	H	1.235	0.05	1
125	16	ILE	HG2	H	0.953	0.05	1
126	16	ILE	H	H	7.899	0.05	1
127	16	ILE	N	N	119.135	0.5	1
128	17	LEU	CA	C	57.363	0.5	1
129	17	LEU	CB	C	41.928	0.5	1
130	17	LEU	HA	H	4.104	0.05	1
131	17	LEU	HB2	H	1.759	0.05	1
132	17	LEU	HD1	H	0.941	0.05	1
133	17	LEU	HD2	H	0.914	0.05	1
134	17	LEU	HG	H	1.626	0.05	1
135	17	LEU	H	H	8.107	0.05	1
136	17	LEU	N	N	118.732	0.5	1
137	18	ASP	CA	C	54.922	0.5	1
138	18	ASP	CB	C	37.578	0.5	1
139	18	ASP	HA	H	4.662	0.05	1
140	18	ASP	HB2	H	3.128	0.05	1
141	18	ASP	HB3	H	2.955	0.05	1
142	18	ASP	H	H	8.170	0.05	1
143	18	ASP	N	N	116.426	0.5	1
144	19	ALA	CA	C	55.943	0.5	1
145	19	ALA	CB	C	26.796	0.5	1
146	19	ALA	HA	H	4.291	0.05	1
147	19	ALA	HB	H	1.647	0.05	1
148	19	ALA	H	H	8.651	0.05	1
149	19	ALA	N	N	123.113	0.5	1
150	20	GLY	CA	C	46.379	0.5	1
151	20	GLY	HA2	H	3.873	0.05	1
152	20	GLY	HA3	H	3.776	0.05	1
153	20	GLY	H	H	8.213	0.05	1
154	20	GLY	N	N	103.144	0.5	1
155	21	TYR	CA	C	56.862	0.5	1
156	21	TYR	CB	C	38.321	0.5	1
157	21	TYR	HA	H	4.266	0.05	1
158	21	TYR	HB2	H	2.828	0.05	1
159	21	TYR	HB3	H	2.822	0.05	1
160	21	TYR	H	H	7.697	0.05	1
161	21	TYR	N	N	117.775	0.5	1
162	22	PHE	CA	C	58.875	0.5	1
163	22	PHE	CB	C	39.510	0.5	1
164	22	PHE	HA	H	4.540	0.05	1
165	22	PHE	HB2	H	3.319	0.05	1
166	22	PHE	HB3	H	3.044	0.05	1
167	22	PHE	H	H	7.760	0.05	1
168	22	PHE	N	N	114.045	0.5	1
169	23	LEU	CA	C	55.118	0.5	1
170	23	LEU	CB	C	42.361	0.5	1
171	23	LEU	HA	H	4.331	0.05	1
172	23	LEU	HB2	H	1.766	0.05	1
173	23	LEU	HD1	H	0.905	0.05	1
174	23	LEU	HD2	H	0.901	0.05	1
175	23	LEU	HG	H	1.621	0.05	1
176	23	LEU	H	H	7.521	0.05	1
177	23	LEU	N	N	116.377	0.5	1
178	24	PRO	CA	C	68.713	0.5	1
179	24	PRO	CB	C	39.183	0.5	1
180	24	PRO	HA	H	4.322	0.05	1
181	24	PRO	HB2	H	2.567	0.05	1
182	24	PRO	HB3	H	3.397	0.05	1
183	24	PRO	HG2	H	2.461	0.05	1
184	25	LEU	CA	C	55.147	0.5	1
185	25	LEU	CB	C	40.498	0.5	1
186	25	LEU	HA	H	4.288	0.05	1
187	25	LEU	HB2	H	1.817	0.05	1
188	25	LEU	HB3	H	1.800	0.05	1
189	25	LEU	HG	H	1.575	0.05	1
190	25	LEU	H	H	7.526	0.05	1
191	25	LEU	N	N	116.123	0.5	1
192	26	ARG	CA	C	57.646	0.5	1
193	26	ARG	CB	C	30.563	0.5	1
194	26	ARG	HA	H	4.336	0.05	1
195	26	ARG	HB2	H	1.812	0.05	1
196	26	ARG	HB3	H	1.929	0.05	1
197	26	ARG	HD2	H	3.178	0.05	1
198	26	ARG	HD3	H	3.178	0.05	1
199	26	ARG	HE	H	7.322	0.05	1
200	26	ARG	HG2	H	1.672	0.05	1
201	26	ARG	HG3	H	1.648	0.05	1
202	26	ARG	H	H	7.858	0.05	1
203	26	ARG	N	N	116.662	0.5	1
204	26	ARG	NE	N	108.582	0.5	1
205	27	HSL	CB	C	30.146	0.5	1
206	27	HSL	HA	H	4.532	0.05	1
207	27	HSL	HB2	H	2.390	0.05	1
208	27	HSL	HB3	H	2.571	0.05	1
209	27	HSL	HG2	H	3.841	0.05	1
210	27	HSL	HG3	H	3.841	0.05	1
211	27	HSL	H	H	8.219	0.05	1
212	27	HSL	N	N	112.928	0.5	1

Appendix C-2. **Procheck analysis of TMIV of hNHE1**

The following pages contain analysis of the NMR ensemble structure of TMIV of hNHE1, as per the pdb coordinates **1Y4E** from the protein data bank. Note that I have only analyzed the 60 lowest energy structures from the 100 structures that were deposited in the pdb. Complete Ramachandran plot analysis for TMIV and individual residues are reported, as per procheck-NMR.

Ramachandran Plot

1Y4E (60 models)



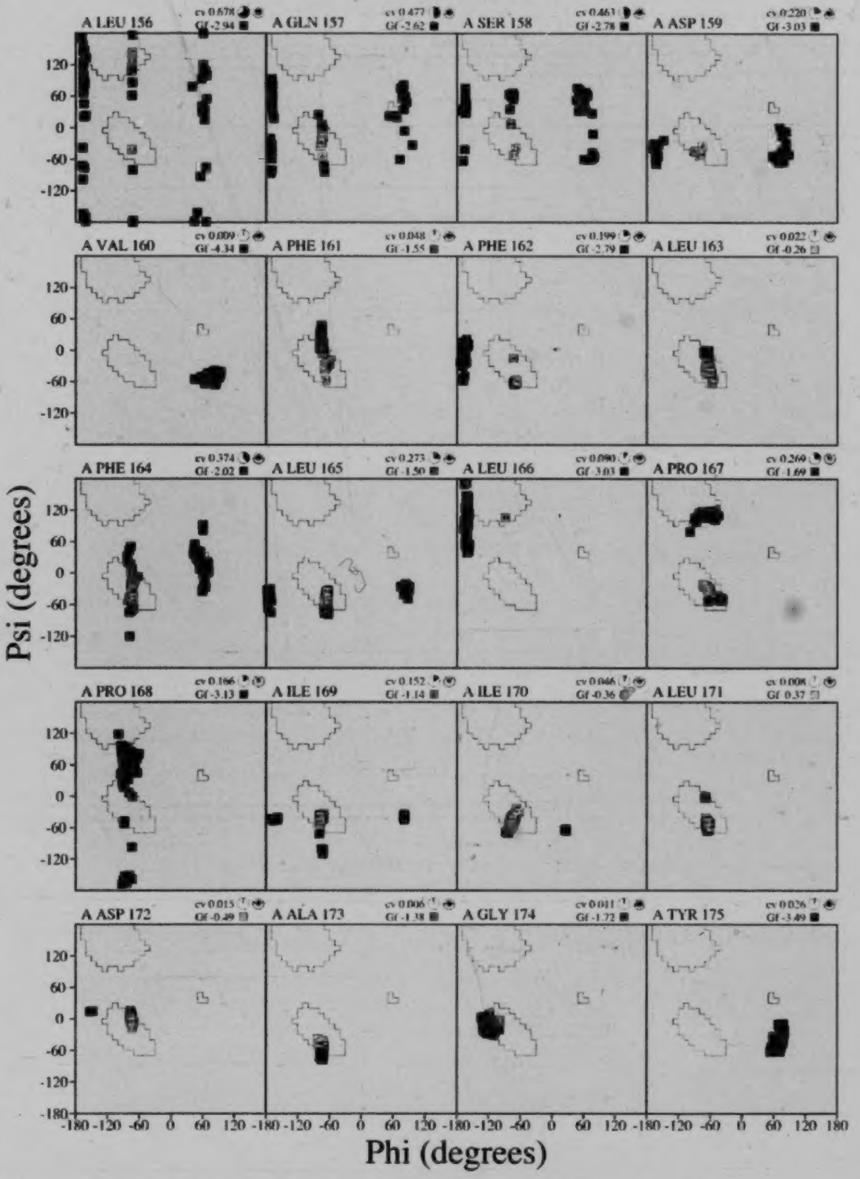
Plot statistics

Residues in most favoured regions [A,B,L]	549	43.6%
Residues in additional allowed regions [a,b,l,p]	403	32.0%
Residues in generously allowed regions [~a,~b,~l,~p]	149	11.8%
Residues in disallowed regions	159	12.6%
	----	------
Number of non-glycine and non-proline residues	1260	100.0%
Number of end-residues (excl. Gly and Pro)	120	
Number of glycine residues (shown as triangles)	60	
Number of proline residues	180	

Total number of residues	1620	

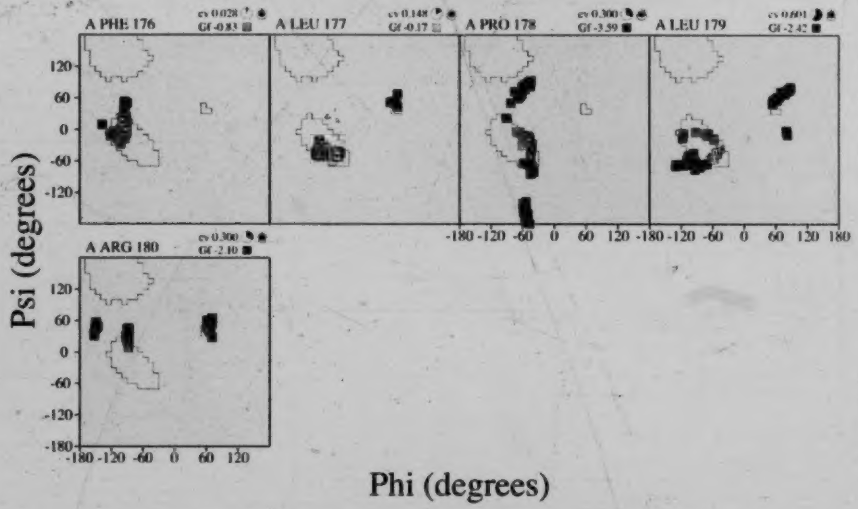
Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms
and R-factor no greater than 20%, a good quality model would be expected
to have over 90% in the most favoured regions.
Model numbers shown inside each data point.

Ensemble Ramachandran plots
1Y4E (60 models)



cv = Circular Variance (low values signify high clustering of the data points). ● Accessible ⓘ Buried
Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high res. Xstal structures
Data points coloured according to G-factor: Unfavourable

Ensemble Ramachandran plots
1Y4E (60 models)



cv = Circular Variance (low values signify high clustering of the data points). ● Accessible ◑ Buried

Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high-res. Xstal structures

Data points coloured according to G-factor: Unfavourable

APPENDIX D

The inhibitory region of cardiac troponin I: structure and function - chemical shift and procheck structural analysis

This appendix contains supplementary information regarding structural characterization of the presented NMR solution structure of the C-lobe of cardiac troponin C in complex with the inhibitory region of cardiac troponin I (128-147), described in Chapter III of this thesis. This work was published in the July 2003 issue of the *Journal of Biological* [*J Biol Chem.* 2003 Jul 18;**278(29)**:27024-34.]. All chemical shift data are reported as nmrView standard output and do not reflect the precision of the chemical shifts. All aromatic ^{13}C chemical shifts reported for the C-lobe are reported as a folded dimensional value.

Appendix D-1. **Assigned ^{1}H, ^{13}C and ^{15}N chemical shifts of unbound cIp in solution**

```
128.N      119.690 1      131.HA       4.099 1      134.HG       1.661 1
128.HT1      8.144 1      131.CB      38.764 1      134.CD1     25.099 1
128.CA      62.324 1      131.HB       1.777 1      134.HD11     0.947 2
128.HA       4.277 1      131.CG1     27.251 1      134.CD2     23.537 2
128.CB      69.811 1      131.HG12     1.361 1      134.HD21     0.880 2
128.HB       4.206 1      131.HG11     1.113 1      135.N      119.945 1
128.CG2     21.712 1      131.CD1     12.785 1      135.HN       8.257 1
128.HG21     1.222 1      131.HD11     0.811 1      135.CA      56.499 1
129.N      122.735 1      131.CG2     17.525 1      135.HA       4.267 1
129.HN       8.430 1      131.HG21     0.802 2      135.CB      30.431 1
129.CA      56.076 1      132.N      123.955 1      135.HB2      1.893 2
129.HA       4.337 1      132.HN       8.227 1      135.HB1      1.851 2
129.CB      29.453 1      132.CA      57.779 1      135.CG      27.259 1
129.HB2      2.076 2      132.HA       4.614 1      135.HG2      1.694 2
129.HB1      1.988 2      132.CB      39.857 1      135.HG1      1.645 2
129.CG      33.835 1      132.HB2      3.117 2      135.CD      43.446 1
129.HG2      2.356 2      132.HB1      2.984 2      135.HD2      3.194 2
130.N      123.003 1      133.N      122.136 1      136.N      108.922 1
130.HN       8.315 1      133.HN       8.221 1      136.HN       8.241 1
130.CA      56.839 1      133.CA      54.079 1      136.CA      45.661 1
130.HA       4.269 1      133.HA       4.614 1      136.HA2      3.908 2
130.CB      32.969 1      133.CB      41.414 1      137.N      120.614 1
130.HB2      1.705 2      133.HB2      2.708 2      137.HN       8.034 1
130.CG      24.836 1      133.HB1      2.618 2      137.CA      56.547 1
130.HG2      1.391 2      134.N      123.446 1      137.HA       4.213 1
130.HG1      1.318 2      134.HN       8.210 1      137.CB      32.834 1
130.CE      42.214 1      134.CA      55.821 1      137.HB2      1.653 2
130.HE2      2.965 2      134.HA       4.273 1      137.CG      24.652 1
130.HE1      2.965 2      134.CB      42.161 1      137.HG2      1.269 2
131.N      121.461 1      134.HB2      1.661 2      137.HG1      1.209 2
131.HN       7.962 1      134.HB1      1.661 2      137.HD2      2.938 2
131.CA      61.179 1      134.CG      29.101 1      137.CE      42.214 1
```

137.HE2	2.971	2
137.HE1	2.971	2
138.N	120.085	1
138.HN	8.089	1
138.CA	57.523	1
138.HA	4.637	1
138.CB	39.605	1
138.HB2	3.161	2
138.HB1	3.006	2
139.N	123.315	1
139.HN	8.101	1
139.CA	56.076	1
139.HA	4.289	1
139.CB	33.216	1
139.HB2	1.759	2
139.HB1	1.687	2
139.CG	24.758	1
139.HG2	1.384	2
139.HG1	1.361	2
139.CE	42.214	1
139.HE2	2.971	2
139.HE1	2.971	2
140.N	123.814	1
140.HN	8.270	1
140.CA	54.163	1
140.HA	4.552	1
140.CB	30.250	1
140.HB2	1.857	2
140.HB1	1.736	2
140.CG	27.035	1
140.HG2	1.705	2
140.HG1	1.705	2
140.CD	43.446	1
140.HD2	3.228	2
141.CA	63.269	1
141.HA	4.488	1
141.CB	32.220	1
141.HB2	2.316	2
141.HB1	1.911	2
141.CG	27.444	1
141.HG2	2.042	2
141.HG1	2.042	2
141.CD	50.785	1
141.HD2	3.835	2
141.HD1	3.684	2
142.N	115.395	1
142.HN	8.238	1
142.CA	62.051	1
142.HA	4.287	1
142.CB	69.887	1
142.HB	4.165	1
142.CG2	21.712	1
142.HG21	1.214	1
143.N	125.383	1
143.HN	8.231	1
143.CA	55.214	1
143.HA	4.378	1
143.CB	42.529	1
143.HB2	1.603	2
143.HB1	1.603	2
143.CG	29.038	1
143.HG	1.595	1
143.CD1	24.858	1
143.HD11	0.925	2
143.CD2	23.718	2
143.HD21	0.870	2
144.N	122.749	1
144.HN	8.304	1
144.CA	56.076	1
144.HA	4.331	1
144.CB	30.913	1
144.HB2	1.817	2
144.HB1	1.754	2
144.CG	27.213	1
144.HG2	1.658	2
144.HG1	1.582	2
144.CD	43.446	1
144.HD2	3.201	2
145.N	123.308	1
145.HN	8.353	1
145.CA	56.076	1
145.HA	4.374	1
145.CB	30.913	1
145.HB2	1.823	2
145.HB1	1.761	2
145.CG	27.213	1
145.HG2	1.644	2
145.HG1	1.590	2
145.CD	43.446	1
145.HD2	3.206	2
146.N	121.942	1
146.HN	8.196	1
146.CA	62.297	1
146.HA	4.135	1
146.CB	32.980	1
146.HB	2.088	1
146.CG2	20.473	1
146.HG21	0.941	2
146.CG1	21.175	2
146.HG11	0.941	2
147.N	125.256	1
147.HN	8.392	1
147.CA	56.076	1
147.HA	4.302	1
147.CB	30.925	1
147.HB2	1.853	2
147.HB1	1.782	2
147.CG	27.236	1
147.HG2	1.674	2
147.HG1	1.625	2
147.CD	43.446	1
147.HD2	3.220	2

Appendix D-2. ^{1}H, ^{13}C and ^{15}N chemical shifts of cCTnC•2Ca2$^+$•cIp binary complex

Appendix D-2A. C-lobe assigned chemical shifts

89.HA	4.122	1
89.CE	16.955	1
89.HE1	2.118	1
92.N	120.678	1
92.HN	8.248	1
92.CA	56.488	1
92.HA	4.433	1
92.CB	33.521	1
92.HB2	1.773	2
92.HB1	1.861	2
92.CG	20.785	1
92.HG2	1.499	2
92.HG1	1.666	2
92.CD	25.260	1
92.HD2	1.499	2
92.HD1	1.666	2
92.CE	42.292	1
92.HE2	3.006	2
92.HE1	3.006	2
93.HN	8.661	1
93.CA	58.673	1
93.HA	4.481	1
93.CB	64.351	1
93.HB2	4.033	2
93.HB1	4.208	2
94.N	122.275	1
94.HN	8.760	1
94.CA	59.496	1
94.HA	4.083	1
94.CB	29.452	1
94.HB2	2.116	2
94.HB1	0.000	2
94.HG2	2.398	2
95.N	120.101	1
95.HN	8.535	1
95.CA	59.525	1
95.HA	4.133	1
95.CB	29.529	1
95.HB2	2.092	2

95.HB1	0.000	2
95.CG	36.693	1
95.HG2	2.358	2
96.N	119.871	1
96.HN	8.040	1
96.CA	59.146	1
96.HA	4.247	1
96.CB	30.202	1
96.HB2	2.124	2
96.HB1	2.264	2
96.CG	37.109	1
96.HG2	2.353	2
96.HG1	2.453	2
97.N	120.160	1
97.HN	8.363	1
97.CA	58.114	1
97.HA	4.366	1
97.CB	41.882	1
97.HB2	1.575	2
97.HB1	2.144	2
97.CG	27.125	1
97.HG	1.874	1
97.CD1	26.245	1
97.HD11	0.846	2
97.CD2	23.635	2
97.HD21	0.846	2
98.N	116.209	1
98.HN	8.623	1
98.CA	61.784	1
98.HA	4.356	1
98.CB	62.541	1
98.HB2	4.114	2
98.HB1	4.357	2
99.N	121.718	1
99.HN	7.948	1
99.CA	57.496	1
99.HA	4.482	1
99.CB	40.569	1
99.HB2	2.690	2
99.HB1	2.907	2
100.N	121.531	1
100.HN	8.045	1
100.CA	58.160	1
100.HA	4.171	1
100.CB	42.697	1
100.HB2	2.034	2
100.HB1	2.119	2
100.HG	1.733	1
100.CD1	25.260	1
100.HD11	1.114	2
100.CD2	23.801	1
100.HD21	0.827	2
101.N	119.632	1
101.HN	8.478	1
101.CA	62.390	1
101.HA	4.350	1
101.CB	39.332	1
101.HB2	3.158	2
101.HB1	2.838	2
101.CD1	59.450	3
101.HD1	6.434	3
101.CE1	59.450	3
101.HE1	6.952	3
101.CE2	59.450	1
101.HE2	6.952	3
101.CD2	59.450	3
101.HD2	6.434	3
102.N	114.877	1
102.HN	7.914	1
102.CA	59.021	1
102.HA	3.947	1
102.CB	28.000	1
102.HB2	1.744	2
102.HB1	1.951	2
102.CG	25.260	1
102.HG2	0.963	2
102.HG1	1.714	2
102.CD	43.405	1
102.HD2	3.271	2
102.HD1	3.271	2
103.N	117.276	1
103.HN	7.725	1
103.CA	58.233	1
103.HA	3.915	1
103.CB	32.620	1
103.HB2	1.860	2
103.HB1	2.040	2
103.CG	31.702	1
103.HG2	2.276	2
103.HG1	2.038	2
103.CE	16.918	1
103.HE1	1.941	1
104.N	114.765	1
104.HN	7.363	1
104.CA	59.588	1
104.HA	4.436	1
104.CB	39.885	1
104.HB2	2.555	2
104.HB1	3.005	2
104.CD1	59.716	3
104.HD1	7.429	3
104.CE1	59.594	3
104.HE1	7.296	3
104.CE2	59.594	1
104.HE2	7.296	3
104.CD2	59.716	3
104.HD2	7.429	3
105.N	120.137	1
105.HN	7.714	1
105.CA	52.426	1
105.HA	4.523	1
105.CB	38.028	1
105.HB2	1.573	2
105.HB1	2.471	2
106.N	125.475	1
106.HN	7.537	1
106.CA	58.578	1
106.HA	3.976	1
106.CB	32.801	1
106.HB2	1.919	2
106.HB1	1.861	2
106.HG2	1.616	2
106.HG1	1.579	2
106.CD	28.482	1
106.HD2	1.770	2
106.CE	42.414	1
106.HE2	3.096	2
106.HE1	3.096	2
107.N	113.453	1
107.HN	7.951	1
107.CA	51.696	1
107.HA	4.729	1
107.CB	36.788	1
107.HB2	2.859	2
107.HB1	3.307	2
107.ND2	112.480	1
107.HD21	6.696	2
107.HD22	7.874	2
108.N	121.733	1
108.HN	7.840	1
108.CA	53.194	1
108.HA	4.025	1
108.CB	16.973	1
108.HB1	1.366	1
109.N	117.927	1
109.HN	8.468	1
109.CA	53.173	1
109.HA	4.688	1
109.CB	41.047	1
109.HB2	2.469	2
109.HB1	3.210	2
110.N	112.084	1
110.HN	10.360	1
110.CA	45.068	1
110.HA2	3.453	2
110.HA1	4.047	2
111.N	115.946	1
111.HN	8.016	1
111.CA	56.061	1
111.HA	5.278	1
111.CB	43.044	1
111.HB2	2.557	2
111.HB1	2.721	2
111.CD1	62.031	3
111.HD1	6.740	3
111.CE1	62.031	3
111.HE1	6.853	3
111.CE2	62.031	1
111.HE2	6.853	3
111.CD2	62.031	3
111.HD2	6.740	3
112.N	126.098	1
112.HN	9.955	1
112.CA	60.067	1
112.HA	4.853	1
112.CB	38.492	1
112.HB	1.818	1
112.CG1	27.125	1
112.HG12	0.482	1
112.HG11	0.962	1
112.CD1	13.327	1
112.HD11	0.272	1
112.CG2	17.429	1

112.HG21	0.869	2
113.N	128.604	1
113.HN	8.902	1
113.CA	52.003	1
113.HA	4.871	1
113.CB	41.855	1
113.HB2	2.494	2
113.HB1	3.354	2
114.N	118.335	1
114.HN	8.499	1
114.CA	58.974	1
114.HA	3.942	1
114.CB	42.382	1
114.HB2	1.719	2
114.HB1	1.719	2
114.CG	27.498	1
114.HG	1.724	1
114.CD1	24.514	1
114.HD11	0.963	2
114.CD2	25.126	2
114.HD21	0.954	2
115.N	118.231	1
115.HN	7.881	1
115.CA	57.744	1
115.HA	4.380	1
115.CB	40.729	1
115.HB2	2.675	2
115.HB1	2.809	2
116.N	121.142	1
116.HN	8.558	1
116.CA	58.749	1
116.HA	4.087	1
116.CB	29.951	1
116.HB2	2.333	2
116.HB1	2.441	2
116.HG2	2.843	2
117.N	122.660	1
117.HN	8.939	1
117.CA	57.814	1
117.HA	3.989	1
117.CB	42.083	1
117.HB2	1.535	2
117.HB1	1.956	2
117.CG	26.752	1
117.HG	1.593	1
117.CD1	26.379	1
117.HD11	0.851	2
117.CD2	23.768	2
117.HD21	0.851	2
118.N	118.229	1
118.HN	8.099	1
118.CA	60.834	1
118.HA	3.799	1
118.CB	32.415	1
118.HB2	2.040	2
118.HB1	1.768	2
118.CG	25.633	1
118.HG2	1.340	2
118.HG1	1.609	2
118.CD	26.377	1
118.HD2	1.335	2
118.HD1	1.593	2
118.CE	42.083	1
118.HE2	2.890	2
118.HE1	2.957	2
119.N	118.416	1
119.HN	7.583	1
119.CA	64.611	1
119.HA	3.805	1
119.CB	38.085	1
119.HB	2.042	1
119.CG1	28.989	1
119.HG12	1.793	1
119.HG11	1.318	1
119.CD1	12.954	1
119.HD11	0.937	1
119.CG2	17.802	1
119.HG21	1.096	2
120.N	120.268	1
120.HN	7.820	1
120.CA	58.548	1
120.HA	3.527	1
120.CB	33.051	1
120.HB2	1.939	2
120.HB1	1.574	2
120.CG	31.571	1
120.HG2	1.721	2
120.HG1	1.409	2
120.CE	17.382	1
120.HE1	1.790	1
121.N	117.191	1
121.HN	8.267	1
121.CA	56.986	1
121.HA	4.101	1
121.CB	41.029	1
121.HB2	1.442	2
121.HB1	1.846	2
121.CG	27.498	1
121.HG	1.855	1
121.CD1	22.650	1
121.HD11	0.751	2
121.CD2	26.379	2
121.HD21	0.827	2
122.N	119.397	1
122.HN	8.298	1
122.CA	58.320	1
122.HA	4.237	1
122.CB	28.426	1
122.HB2	2.210	2
122.HB1	2.279	2
122.CG	34.583	1
122.HG2	2.423	2
122.HG1	2.617	2
122.NE2	110.462	1
122.HE21	7.407	2
122.HE22	6.815	2
123.N	122.241	1
123.HN	7.873	1
123.CA	54.079	1
123.HA	4.375	1
123.CB	18.674	1
123.HB1	1.681	1
124.N	107.542	1
124.HN	7.639	1
124.CA	62.876	1
124.HA	4.341	1
124.CB	70.539	1
124.HB	4.425	1
124.CG2	21.904	1
124.HG21	1.315	1
125.N	109.134	1
125.HN	7.900	1
125.CA	45.662	1
125.HA2	3.877	2
125.HA1	4.221	2
126.N	119.884	1
126.HN	7.830	1
126.CA	55.674	1
126.HA	4.434	1
126.CB	30.881	1
126.HB2	1.799	2
126.HB1	2.045	2
126.CG	36.159	1
126.HG2	2.239	2
126.HG1	2.239	2
127.N	117.852	1
127.HN	8.641	1
127.CA	62.787	1
127.HA	4.295	1
127.CB	69.124	1
127.HB	4.153	1
127.CG2	21.531	1
127.HG21	1.199	1
128.N	125.341	1
128.HN	8.174	1
128.CA	59.339	1
128.HA	4.635	1
128.CB	39.914	1
128.HB	1.923	1
128.CG1	27.125	1
128.HG12	1.412	1
128.HG11	1.297	1
128.CD1	13.327	1
128.HD11	0.745	1
128.CG2	18.175	1
128.HG21	0.910	2
129.N	117.515	1
129.HN	9.000	1
129.CA	60.623	1
129.HA	4.639	1
129.CB	71.815	1
129.HB	4.660	1
129.CG2	21.531	1
129.HG21	1.316	1
130.N	120.578	1
130.HN	8.953	1
130.CA	60.177	1
130.HA	3.977	1
130.CB	29.228	1
130.HB2	2.085	2
130.HB1	2.358	2
130.CG	36.703	1
130.HG2	2.381	2

131.N	117.811	1
131.HN	8.152	1
131.CA	57.102	1
131.HA	4.404	1
131.CB	40.523	1
131.HB2	2.520	2
131.HB1	2.691	2
132.N	120.055	1
132.HN	7.661	1
132.CA	57.450	1
132.HA	4.392	1
132.CB	41.118	1
132.HB2	2.509	2
132.HB1	3.055	2
133.N	119.887	1
133.HN	7.756	1
133.CA	64.882	1
133.HA	3.548	1
133.CB	37.470	1
133.HB	1.996	1
133.CG1	28.989	1
133.HG12	1.735	1
133.HG11	1.050	1
133.CD1	13.700	1
133.HD11	0.844	1
133.CG2	18.175	1
133.HG21	0.923	2
134.N	119.271	1
134.HN	7.966	1
134.CA	59.964	1
134.HA	3.885	1
134.CB	29.371	1
134.HB2	2.218	2
134.HB1	2.118	2
134.CG	36.688	1
134.HG2	2.255	2
134.HG1	2.482	2
135.N	118.863	1
135.HN	7.872	1
135.CA	59.309	1
135.HA	4.026	1
135.CB	29.458	1
135.HB2	2.156	2
135.HB1	2.394	2
135.CG	36.129	1
135.HG2	2.375	2
135.HG1	2.417	2
136.N	121.062	1
136.HN	7.929	1
136.CA	58.194	1
136.HA	4.130	1
136.CB	41.931	1
136.HB2	1.583	2
136.HB1	1.892	2
136.CG	27.498	1
136.HG	1.791	1
136.CD1	25.633	1
136.HD11	0.844	2
136.CD2	24.380	2
136.HD21	0.871	2
137.N	117.811	1
137.HN	8.152	1
137.CA	57.691	1
137.HA	4.152	1
137.CB	31.127	1
137.HB2	1.886	2
137.HB1	2.191	2
137.CG	32.345	1
137.HG2	2.441	2
137.HG1	2.578	2
137.CE	17.318	1
137.HE1	1.941	1
138.N	117.533	1
138.HN	7.818	1
138.CA	59.162	1
138.HA	4.067	1
138.CB	32.295	1
138.HB2	1.927	2
138.HB1	1.927	2
138.CG	25.260	1
138.HG2	1.487	2
138.HG1	1.674	2
138.CD	29.228	1
138.HD2	1.718	2
138.HD1	1.718	2
138.CE	42.192	1
138.HE2	2.983	2
138.HE1	2.983	2
139.N	117.361	1
139.HN	7.779	1
139.CA	56.549	1
139.HA	4.506	1
139.CB	41.831	1
139.HB2	2.840	2
139.HB1	2.872	2
140.N	104.406	1
140.HN	8.125	1
140.CA	45.702	1
140.HA2	3.011	2
140.HA1	4.027	2
141.N	120.220	1
141.HN	8.225	1
141.CA	53.249	1
141.HA	4.606	1
141.CB	39.657	1
141.HB2	2.429	2
141.HB1	2.999	2
142.N	126.030	1
142.HN	7.799	1
142.CA	57.788	1
142.HA	4.200	1
142.CB	33.007	1
142.HB2	1.963	2
142.HB1	1.771	2
142.CG	24.514	1
142.HG2	1.551	2
142.HG1	1.704	2
142.CE	42.414	1
142.HE2	3.117	2
142.HE1	3.117	2
143.N	113.803	1
143.HN	8.136	1
143.CA	51.690	1
143.HA	4.848	1
143.CB	37.206	1
143.HB2	2.858	2
143.HB1	3.328	2
143.ND2	112.670	1
143.HD21	6.612	2
143.HD22	7.919	2
144.N	115.724	1
144.HN	7.753	1
144.CA	55.100	1
144.HA	4.499	1
144.CB	37.731	1
144.HB2	2.681	2
144.HB1	3.084	2
144.ND2	111.777	1
144.HD21	6.707	2
144.HD22	7.454	2
145.N	118.121	1
145.HN	8.587	1
145.CA	53.024	1
145.HA	4.741	1
145.CB	40.950	1
145.HB2	2.410	2
145.HB1	3.005	2
146.N	113.028	1
146.HN	10.368	1
146.CA	45.841	1
146.HA2	3.481	2
146.HA1	4.036	2
147.N	115.633	1
147.HN	7.753	1
147.CA	54.050	1
147.HA	4.821	1
147.CB	34.292	1
147.HB2	1.508	2
147.HB1	1.621	2
147.CG	25.700	1
147.HG2	1.260	2
147.HG1	1.316	2
147.CD	43.173	1
147.HD2	2.268	2
147.HD1	2.565	2
148.N	125.883	1
148.HN	9.259	1
148.CA	59.824	1
148.HA	5.228	1
148.CB	39.641	1
148.HB	2.034	1
148.CG1	28.243	1
148.HG12	1.593	1
148.HG11	0.803	1
148.CD1	17.802	1
148.HD11	1.181	1
148.CG2	12.830	1
148.HG21	0.597	2
149.N	128.786	1
149.HN	9.296	1
149.CA	52.432	1
149.HA	5.266	1
149.CB	41.854	1

149.HB2	3.258	2
149.HB1	2.878	2
150.N	118.786	1
150.HN	8.454	1
150.CA	62.170	1
150.HA	5.253	1
150.CB	38.241	1
150.HB2	2.493	2
150.HB1	2.177	2
150.CD1	60.911	3
150.HD1	6.384	3
150.CE1	60.911	3
150.HE1	6.607	3
150.CE2	60.911	1
150.HE2	6.607	3
150.CD2	60.911	3
150.HD2	6.384	3
151.N	117.798	1
151.HN	7.766	1
151.CA	57.970	1
151.HA	4.176	1
151.CB	40.622	1
151.HB2	2.619	2
151.HB1	2.731	2
152.N	120.849	1
152.HN	8.863	1
152.CA	58.760	1
152.HA	4.086	1
152.CB	29.985	1
152.HB2	2.295	2
152.HB1	2.608	2
152.CG	37.193	1
152.HG2	3.063	2
152.HG1	2.877	2
153.N	124.242	1
153.HN	9.251	1
153.CA	61.886	1
153.HA	4.054	1
153.CB	39.797	1
153.HB2	3.298	2
153.HB1	3.371	2
153.CD1	59.978	3
153.HD1	7.069	3
153.CE1	59.970	3
153.HE1	7.463	3
153.CE2	59.970	1
153.HE2	7.463	3
153.CD2	59.978	3
153.HD2	7.069	3
154.N	118.536	1
154.HN	8.192	1
154.CA	58.344	1
154.HA	3.638	1
154.CB	41.645	1
154.HB2	1.326	2
154.HB1	1.734	2
154.CG	26.379	1
154.HG	1.333	1
154.CD1	23.311	1
154.HD11	0.752	2
154.CD2	25.239	2
154.HD21	0.691	2
155.N	116.598	1
155.HN	7.279	1
155.CA	59.056	1
155.HA	4.001	1
155.CB	29.668	1
155.HB2	2.085	2
155.HB1	2.085	2
155.CG	36.073	1
155.HG2	2.342	2
155.HG1	2.371	2
156.N	120.818	1
156.HN	8.348	1
156.CA	59.737	1
156.HA	4.429	1
156.CB	39.888	1
156.HB2	3.131	2
156.HB1	3.343	2
156.CD1	59.508	3
156.HD1	7.015	3
156.CE1	59.429	3
156.HE1	7.172	3
156.CE2	59.429	1
156.HE2	7.172	3
156.CD2	59.508	3
156.HD2	7.015	3
157.N	114.972	1
157.HN	7.975	1
157.CA	54.418	1
157.HA	4.213	1
157.CB	31.325	1
157.HB2	1.624	2
157.HB1	1.739	2
157.CG	31.972	1
157.HG2	1.542	2
157.HG1	1.770	2
157.CE	16.229	1
157.HE1	1.749	1
157.C	0.000	1
158.N	120.035	1
158.HN	7.497	1
158.CA	58.485	1
158.HA	4.074	1
158.CB	32.560	1
158.HB2	1.858	2
158.HB1	1.852	2
158.CG	25.260	1
158.HG2	1.438	2
158.HG1	1.582	2
158.CD	29.362	1
158.HD2	1.705	2
158.CE	42.041	1
158.HE2	2.892	2
158.HE1	2.957	2
159.N	107.858	1
159.HN	7.976	1
159.CA	45.374	1
159.HA2	3.839	2
159.HA1	3.975	2
160.N	118.889	1
160.HN	7.528	1
160.CA	62.301	1
160.HA	4.069	1
160.CB	32.558	1
160.HB	2.050	1
160.CG2	20.785	1
160.HG21	0.785	2
160.CG1	21.531	2
160.HG11	0.819	2
161.N	129.524	1
161.HN	7.914	1
161.CA	58.138	1
161.HA	4.074	1
161.CB	31.371	1
161.HB2	1.872	2
161.HB1	2.034	2
161.CG	36.820	1
161.HG2	2.190	2
161.HG1	2.190	2

Appendix D-2B. clp assigned chemical shifts

128.HN	8.136	1
128.CA	61.738	1
128.HA	3.678	1
128.CB	70.021	1
128.HB	4.118	1
128.CG2	21.571	1
128.HG21	1.278	1
129.N	121.708	1
129.HN	8.288	1
129.CA	56.141	1
129.HA	4.403	1
129.CB	29.740	1
129.HB2	2.088	2
129.HB1	2.007	2
129.CG	33.877	1
129.HG2	2.396	2
129.NE2	112.440	1
129.HE21	7.586	2
129.HE22	6.869	2
130.N	122.964	1
130.HN	8.508	1
130.CA	56.448	1
130.HA	4.269	1
130.CB	32.880	1
130.HB2	1.737	2
130.HB1	1.634	2
130.HG2	1.409	2
130.HG1	1.352	2
130.CE	42.152	1
130.HE2	2.967	2

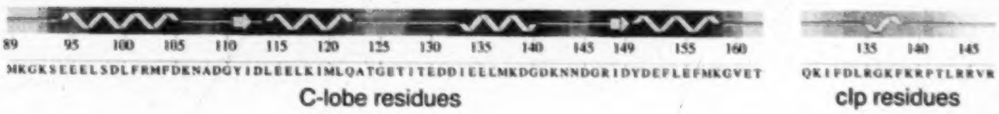
131.N	121.066	1
131.HN	8.135	1
131.CA	61.455	1
131.HA	4.084	1
131.CB	38.659	1
131.HB	1.790	1
131.CG1	27.141	2
131.HG12	1.397	1
131.HG11	1.152	1
131.CD1	12.959	1
131.HD11	0.800	1
131.CG2	17.753	1
131.HG21	0.789	2
132.N	122.677	1
132.HN	8.078	1
132.CA	57.988	1
132.HA	4.582	1
132.CB	39.866	1
132.HB2	3.099	2
132.HB1	2.984	2
132.CD1	59.958	3
132.HD1	7.226	3
132.HE1	7.226	3
132.HE2	7.226	3
132.HD2	7.226	3
133.N	121.767	1
133.HN	8.245	1
133.CA	54.227	1
133.HA	4.667	1
133.CB	41.382	1
133.HB2	2.784	2
133.HB1	2.649	2
134.N	122.923	1
134.HN	8.256	1
134.CA	55.857	1
134.HA	4.234	1
134.CB	42.539	1
134.HB2	1.684	2
134.HB1	1.627	2
134.CG	27.147	1
134.HG	1.606	1
134.CD1	23.907	1
134.HD11	0.851	2
134.CD2	25.355	1
134.HD21	0.957	2
135.N	118.781	1
135.HN	8.167	1
135.CA	57.983	1
135.HA	4.082	1
135.CB	30.156	1
135.HB2	0.000	2
135.HB1	1.877	2
135.CG	27.209	1
135.HG2	1.685	2
135.HG1	1.606	2
135.CD	43.499	1
135.HD1	3.159	2
136.N	108.419	1
136.HN	8.380	1
136.CA	46.112	1
136.HA2	3.934	2
137.N	119.250	1
137.HN	7.924	1
137.CA	56.554	1
137.HA	4.247	1
137.CB	32.851	1
137.HB2	1.646	2
137.HB1	1.546	2
137.CG	24.928	1
137.HG2	1.284	2
137.HG1	1.227	2
137.CD	28.973	1
137.HD2	1.561	2
137.CE	42.134	1
137.HE2	2.864	2
138.N	117.976	1
138.HN	7.924	1
138.CA	57.604	1
138.HA	4.662	1
138.CB	39.312	1
138.HB2	3.183	2
138.HB1	3.001	2
138.HD1	7.266	3
138.HE1	7.266	3
138.CE2	7.266	3
138.HD2	7.266	3
139.N	121.659	1
139.HN	7.995	1
139.CA	55.970	1
139.HA	4.320	1
139.CB	33.303	1
139.HB2	1.703	2
139.HB1	1.806	2
139.CG	24.876	1
139.HG2	1.390	2
139.HG1	1.361	2
139.CD	29.064	1
139.HD2	1.665	2
139.CE	42.193	1
139.HE2	3.000	2
140.N	122.631	1
140.HN	8.283	1
140.CA	54.528	1
140.HA	4.547	1
140.CB	30.162	1
140.HB2	1.888	2
140.HB1	1.792	2
140.CG	27.389	1
140.HG2	1.688	2
140.CD	43.625	1
140.HD2	3.238	2
141.CA	63.256	1
141.HA	4.510	1
141.CB	32.028	1
141.HB2	2.248	2
141.HB1	1.898	2
141.CG	27.509	1
141.HG2	2.036	2
141.HG1	2.036	2
141.CD	50.633	1
141.HD2	3.824	2
141.HD1	3.650	2
142.N	114.181	1
142.HN	8.139	1
142.CA	62.049	1
142.HA	4.271	1
142.CB	69.746	1
142.HB	4.181	1
142.CG2	21.837	1
142.HG21	1.223	1
143.N	124.379	1
143.HN	8.178	1
143.CA	55.365	1
143.HA	4.361	1
143.CB	42.473	1
143.HB2	1.637	2
143.HB1	1.594	2
143.CG	27.211	1
143.HG	1.653	1
143.CD1	23.791	1
143.HD11	0.906	2
143.CD2	24.961	1
143.HD21	0.955	2
144.N	122.023	1
144.HN	8.267	1
144.CA	55.998	1
144.HA	4.289	1
144.HB2	1.832	2
144.HB1	1.761	2
144.HG2	1.649	2
144.HG1	1.582	2
144.HD2	3.188	2
145.N	121.549	1
145.HN	8.288	1
145.CA	55.998	1
145.HA	4.318	1
145.CB	30.913	1
145.HB2	1.837	2
145.HB1	1.761	2
145.HG2	1.639	2
145.HG1	1.587	2
145.HD2	3.190	2
146.N	121.496	1
146.HN	8.171	1
146.CA	62.392	1
146.HA	4.157	1
146.CB	32.881	1
146.HB	2.119	1
146.CG2	21.536	1
146.HG21	0.981	2
146.CG1	20.882	1
146.HG11	0.987	2
147.N	125.248	1
147.HN	8.347	1
147.CA	55.820	1
147.HA	4.385	1
147.CB	31.068	1
147.HB2	1.850	2
147.HB1	1.770	2
147.HG2	1.659	2
147.HG1	1.615	2
147.CD	0.000	1
147.HD2	3.203	2

Appendix D-3. **Procheck analysis of cCTnC•2Ca2⁺•cIp binary complex**

The following pages contain a procheck-NMR analysis of the ensemble structure of the binary complex cCTnC•2Ca2⁺•cIp (30 structures), as per the pdb coordinates **1OZS** from the protein data bank. Within the pdb file, the C-lobe is defined as fragment A and cIp as fragment B. Complete Ramachandran plot analysis for the whole complex, and individual residues are reported, along with an example of predicted secondary structural characteristics as per procheck-NMR.

Model secondary structure

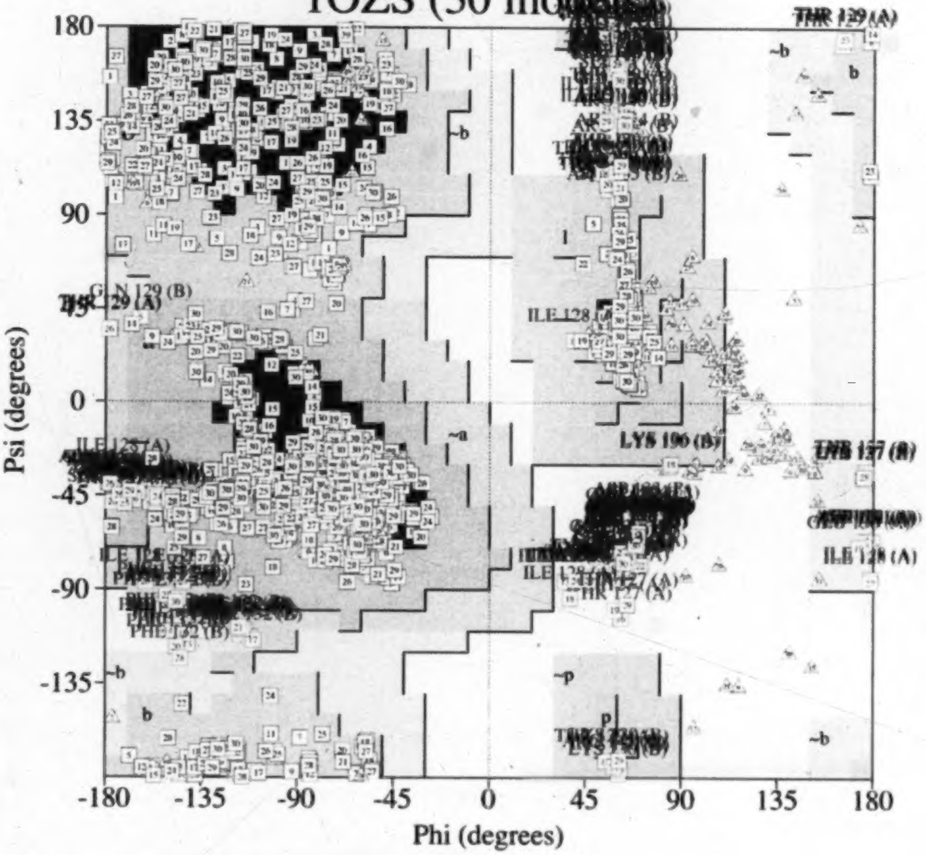
1OZS


89
95
100
105
110
115
120
125
130
135
140
145
149
155
160
MKGKSEEELSDLFRMFDKNADGYIDLEELKIMLQATGETITEDDIEELMKDGDKNNDGRIDYDEFLEFMKGVET
C-lobe residues
135
140
145
QKIFDLRGKFKRPTLRRVR
clp residues

PROCHECK-NMR

Ramachandran Plot

1OZS (30 models)

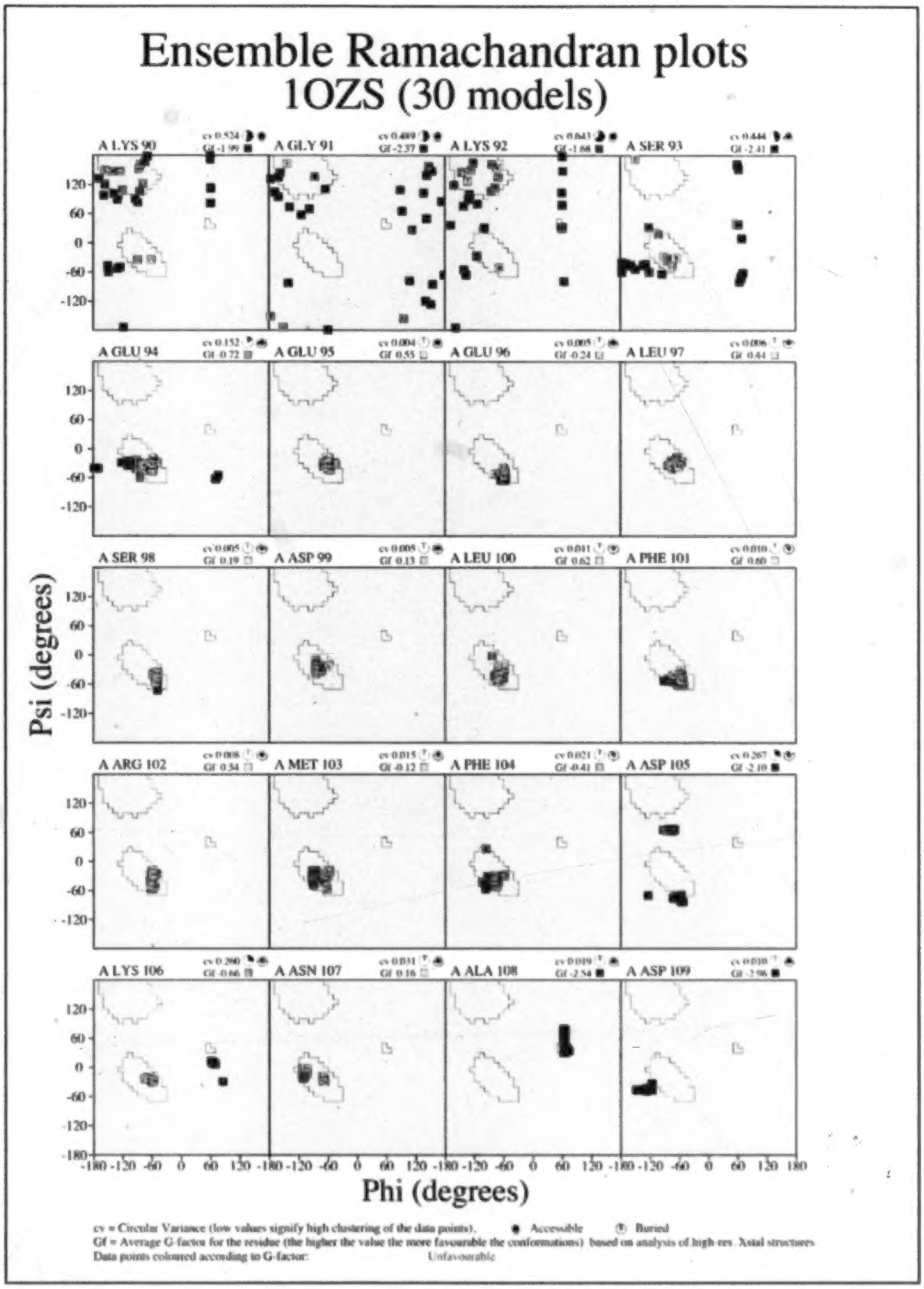


Plot statistics

Residues in most favoured regions [A,B,L]	1677	69.0%
Residues in additional allowed regions [a,b,l,p]	575	23.7%
Residues in generously allowed regions [~a,~b,~l,~p]	95	3.9%
Residues in disallowed regions	83	3.4%
	----	------
Number of non-glycine and non-proline residues	2430	100.0%
Number of end-residues (excl. Gly and Pro)	120	
Number of glycine residues (shown as triangles)	210	
Number of proline residues	30	

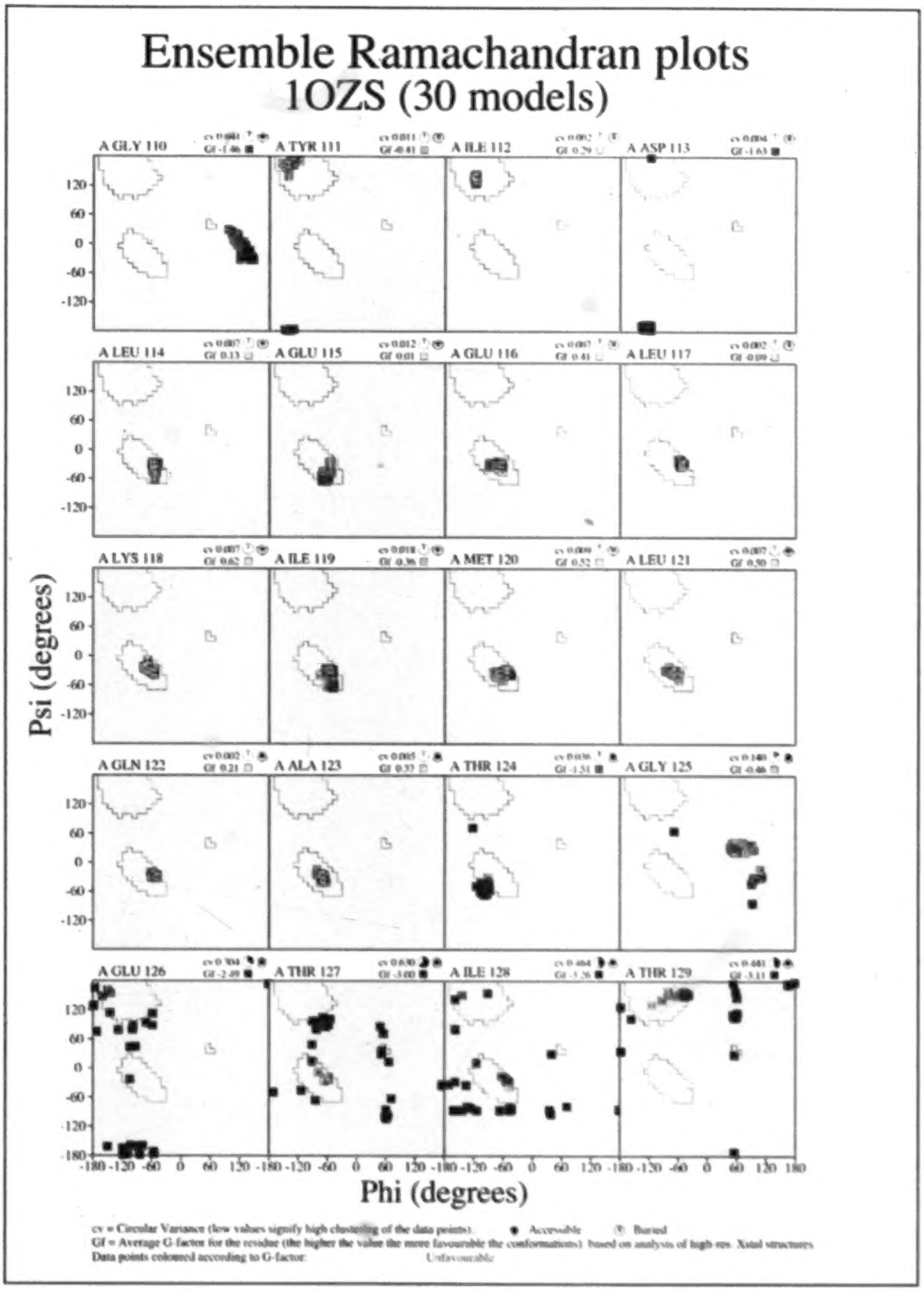
Total number of residues	2790	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.
Model numbers shown inside each data point.



Ensemble Ramachandran plots
1OZS (30 models)

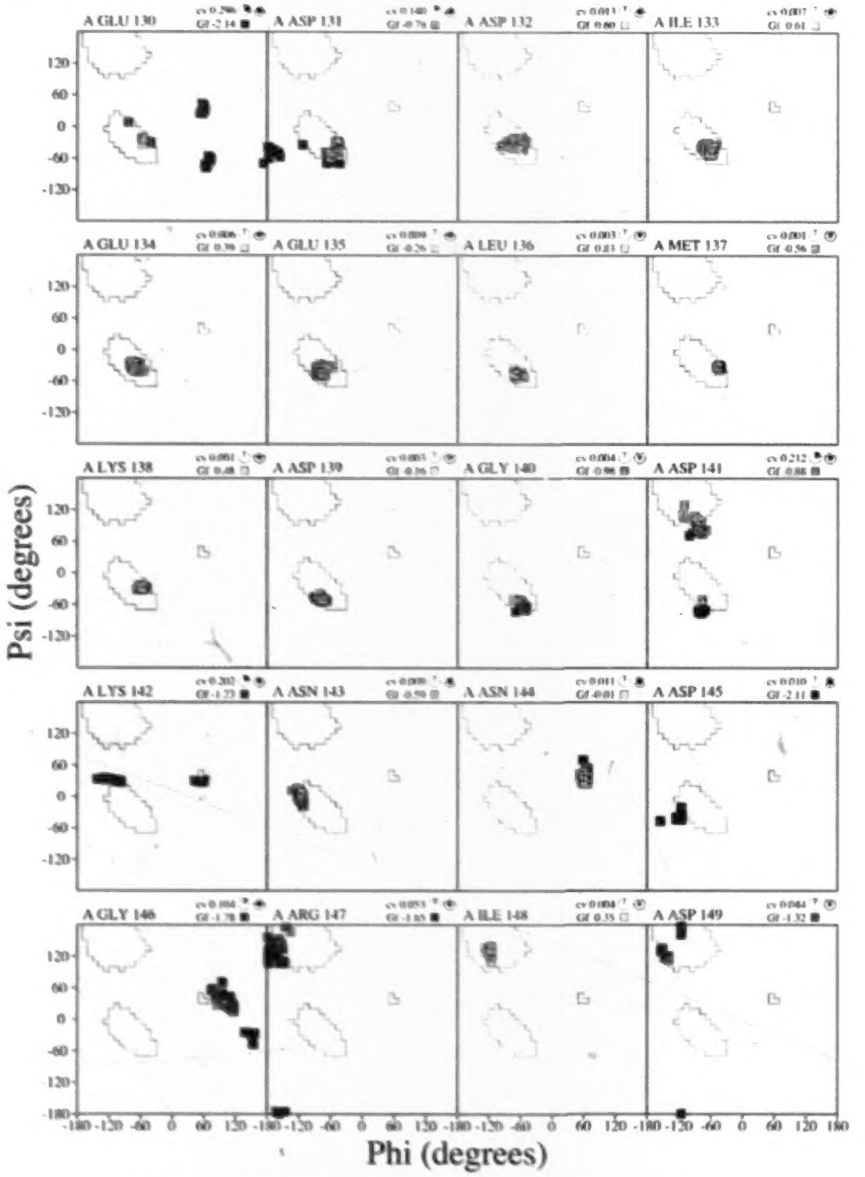
cv = Circular Variance (low values signify high clustering of the data points). ● Accessible ◉ Buried
Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high-res. Xtal structures
Data points coloured according to G-factor: Unfavourable



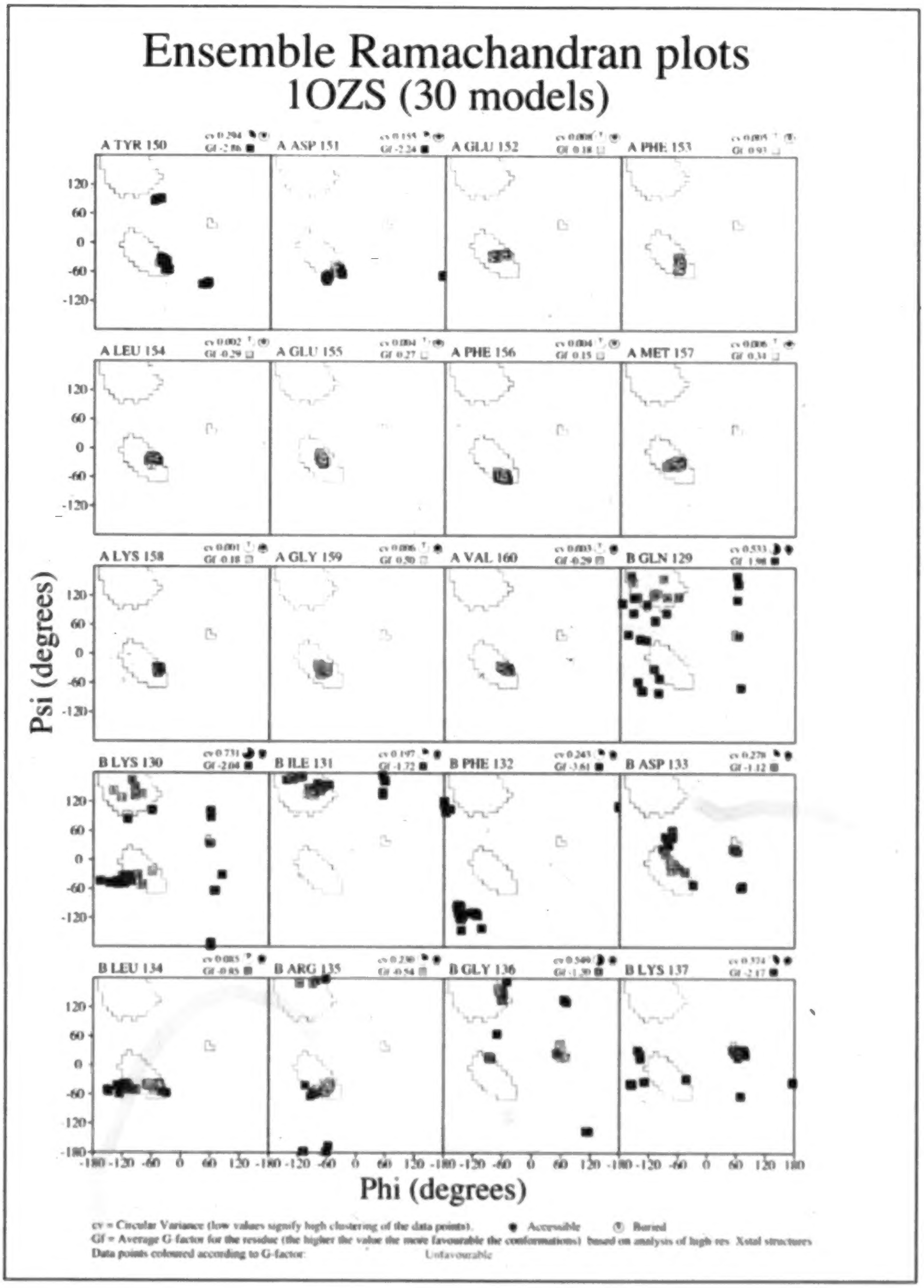
Ensemble Ramachandran plots
1OZS (30 models)

cv = Circular Variance (low values signify high clustering of the data points)   ● Accessible   ⊙ Buried
Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high res. Xstal structures
Data points coloured according to G-factor: Unfavourable

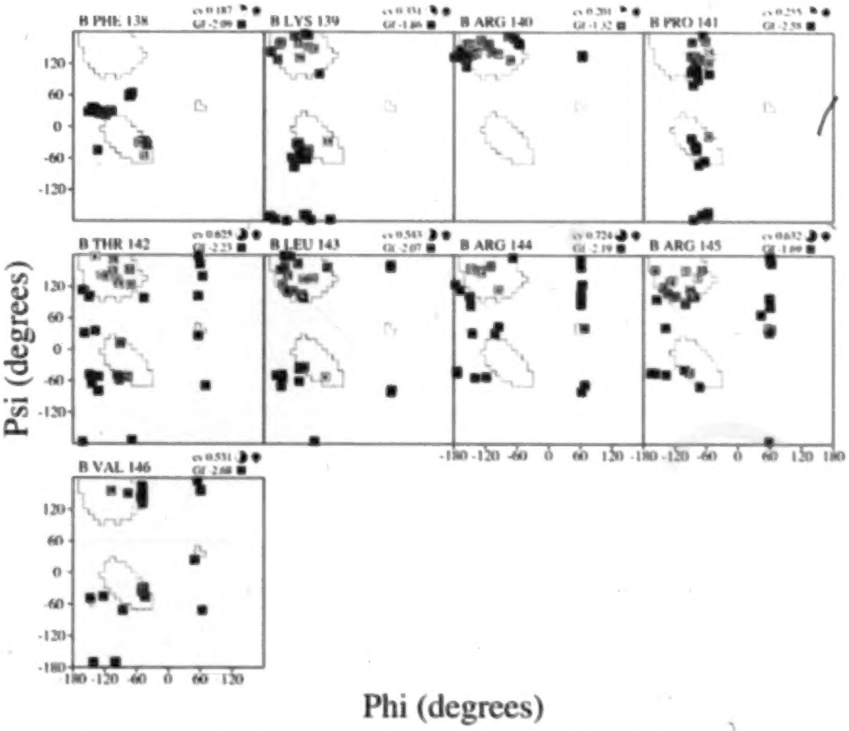
Ensemble Ramachandran plots
1OZS (30 models)



cv = Circular Variance (low values signify high clustering of the data points) ● Accessible ◉ Buried
Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high res. Xtal structures
Data points coloured according to G-factor: Unfavourable


Ensemble Ramachandran plots
1OZS (30 models)
A TYR 150
A ASP 151
A GLU 152
A PHE 153
A LEU 154
A GLU 155
A PHE 156
A MET 157
A LYS 158
A GLY 159
A VAL 160
B GLN 129
B LYS 130
B ILE 131
B PHE 132
B ASP 133
B LEU 134
B ARG 135
B GLY 136
B LYS 137
Psi (degrees)
120
60
0
-60
-120
-180
-180 -120 -60 0 60 120 180
Phi (degrees)
cv = Circular Variance (low values signify high clustering of the data points).
● Accessible
Buried
Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high res. Xstal structures
Data points coloured according to G-factor:
Unfavourable

Ensemble Ramachandran plots
1OZS (30 models)



cv = Circular Variance (low values signify high clustering of the data points) ● Accessible ⊕ Buried
Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high-res. Xtal structures
Data points coloured according to G-factor Unfavourable

APPENDIX E

Using lanthanide ions to align troponin complexes in solution: Order of lanthanide occupancy in cardiac troponin C

This appendix contains supplementary regarding NMR studies of cardiac troponin C in complex with various lanthanide ions, with intent to establish the order of binding occupancy in troponin C. Using the established order of lanthanide binding to the metal binding sites within troponin C, it was possible to obtain partial orientation to measure residual dipolar couplings of the bound inhibitory region to troponin C. The majority of this lanthanide binding data was performed by Grant Lawrence Gay in the Sykes laboratory in the Department of Biochemistry at the University of Alberta, which constituted a large portion of his graduate career (1999-2004). This work was published in the journal *Protein Science* on March 13, 2004 [*Protein Sci.* 2004 Mar;13(3):640-51]. All authors involved in this body of work are Grant L. Gay, Darrin A. Lindhout and Brian D. Sykes. ***The data presented in this appendix is provided only as reviewer reference to supplement the data presented in Chapter III of this thesis.***

OVERVIEW

The potential for using paramagnetic lanthanide ions to partially align troponin C in solution as a tool for the structure determination of bound troponin I peptides has been investigated. A prerequisite for these studies is an understanding of the order of lanthanide ion occupancy in the metal binding sites of the protein. Two-dimensional {^{1}H, ^{15}N}-HSQC NMR spectroscopy has been used to examine the binding order of Ce^{3+}, Tb^{3+}, and Yb^{3+} to both apo- and holo-forms of human cardiac troponin C (cTnC) and of Ce^{3+} to holo-chicken skeletal troponin C (sTnC). The disappearance of cross-peak resonances in the HSQC spectrum was used to determine the order of occupation of the binding sites in both cTnC and sTnC by each lanthanide. For the lanthanides tested, the binding order follows that of the net charge of the binding site residues from most to least negative; the N-lobe calcium binding sites are the first to be filled followed by the C-lobe sites. Given this binding order for lanthanide ions, it was demonstrated that it is possible to create a cTnC species with one lanthanide in the N-lobe site and two Ca^{2+} ions in the C-lobe binding sites. By using the species $cTnC{\bullet}Yb^{3+}{\bullet}2Ca^{2+}$ it was possible to confer partial alignment on a bound human cardiac troponin I (cTnI) peptide. Residual dipolar couplings (RDCs) were measured for the resonances in the bound ^{15}N-labeled $cTnI_{128-147}$ by using two-dimensional {^{1}H, ^{15}N} inphase antiphase (IPAP) NMR spectroscopy.

ABBREVIATIONS

TnC, troponin C; cTnC, cardiac muscle TnC; sTnC, skeletal muscle TnC; TnI, troponin I; cTnI, cardiac muscle TnI; $cTnI_{128-147}$, cTnI peptide residues 128–147; HSQC, heteronuclear single quantum coherence; IPAP, inphase antiphase HSQC

INTRODUCTION

The troponin complex is a group of proteins that are collectively responsible for the regulation of muscle contraction in striated muscle. Troponin C (TnC) plays a key role in this regulation, as it is the molecule that translates the influx of calcium to the sarcoplasmic reticulum into a series of conformational changes in the thin filament proteins that are necessary for contraction. This cascade of conformational changes, which flows from TnC through troponin I (TnI), troponin T, and tropomyosin to actin, where the activation/inhibition of the actomyosin ATPase occurs and triggers muscle contraction (for reviews, see (*1-14*)). Although both skeletal and cardiac muscles behave similarly in this contraction cascade, they do differ in the isoforms of the proteins involved. With respect to TnC, there are two distinct isoforms, one cardiac and one skeletal. They are both dumbbell-shaped molecules that consist of N- and C-terminal globular domains connected by a helical linker, and contain four EF-hand helix–loop–helix folds that are potential Ca^{2+}-binding sites (sites I to IV; for review, see (*15*). Each globular domain contains two Ca^{2+}-binding sites, with the C-lobe sites (III and IV) having a higher affinity for Ca^{2+} and a potential for Mg^{2+} binding, and the N-lobe sites (I and II) are specific for Ca^{2+} and of lower binding affinity. The key difference between the cardiac (cTnC) and skeletal (sTnC) isoforms is a set of key amino acid substitutions that render site I in cTnC unable to bind Ca^{2+} at physiological concentrations

(*16*). In addition to this, it is known that the metal binding properties of the cardiac isoform differ significantly from the skeletal isoform, particularly with respect to Mg^{2+} (*17*). It is believed that sites III and IV are constantly occupied by Ca^{2+} or Mg^{2+} under physiological conditions and that the C-lobe fulfills a more structural role in anchoring TnC to the rest of the thin filament via TnI. The presence or absence of Ca2+ in the N-lobe sites is the regulating factor in the transmission of the signal through the thin filament.

Both the cardiac and skeletal isoforms differ in their site I functionality, which results in differences in the mechanism by which the N-lobe of each bind to their respective TnI counterparts (*18, 19*). In the skeletal isoform, the binding of Ca2+ to the N-lobe results in a conformational change that unveils a hydrophobic pocket (*20*). In cTnC, conformational change does not occur (*21, 22*). However, it has been demonstrated that both isoforms take a similar conformation when bound to their respective TnI regions: residues 115–131 in skeletal and residues 147–163 in cardiac (*23, 24*). This region of TnI is known to bind to the N-lobe of TnC and modulate the interactions between the N terminus of TnI (residues 1–40 in skeletal and 34–71 in cardiac) and the TnI inhibitory domain (residues 96–115 in skeletal and 128–147 in cardiac) with TnC (for reviews, see (*4, 25*)).

Studies of the interaction of cTnC and cTnI by NMR have been limited to cTnI peptides bound to either the N- or C-lobe or to intact cTnC (*26-28*). This is due to the large size of the cTnC•cTnI complex (43 kDa) and a lack of spectral dispersion for residues of TnI, which results in poor-quality spectra from which to extract NOE restraints for structure elucidation. For this reason, other methods such as the residual dipolar coupling (RDC) have been pursued (*29*). RDCs have the advantages that they can be used in higher-molecular-weight complexes and that they provide long-range distance restraints such as the angle that the amide bonds in the protein make with a magnetic susceptibility tensor (χ) principle axis (*30*). These are very useful constraints for elongated structures such as the TnC–TnI complex.

It has been known for quite some time that lanthanides may be substituted for calcium in calcium-binding EF-hand proteins (*31*). This is due to lanthanide metals having a similar radius and ligand preference in the 3+-oxidation state to that of Ca^{2+}, with a higher charge density (*31-33*). Bound paramagnetic lanthanides have been demonstrated to have the ability to induce partial alignment of proteins in high magnetic fields to an extent similar to other orienting methods (*34-36*). These calcium binding proteins have allowed for the use of RDCs as orientation restraints in the calculation and refinement of several solution structures (*37, 38*). Thus, the calcium binding properties of cTnC make the use of lanthanide substitution ideal for the extraction of orientation-based parameters for structure elucidation of the cardiac troponin complex. It has been noted, however, that the order in which lanthanides bind to EF-hand binding sites differs not only with respect to Ca^{2+} binding, but also among the lanthanides themselves (*32*). By using an oriented protein of known structure, it is possible to extract orientation parameters (*36*) and thereby determine the three-dimensional structure of a bound ligand. As a prerequisite for structural studies, it is necessary to understand lanthanide occupancy in TnC. Although others have addressed lanthanide occupancy in the skeletal isoform (*32*), it is unwise to assume that both isoforms behave the same way with respect to metal ion binding. This has been shown in the studies of the magnesium binding abilities of the two isoforms, which demonstrated that magnesium binds 17-fold stronger to sites III and IV in sTnC than to cTnC (*39*). This work details the binding order of several lanthanide ions (Ce^{3+}, Tb^{3+}, and Yb^{3+}) to apo- and Ca^{2+}-saturated cTnC, as well as the binding order of Ce^{3+} to Ca^{2+}-saturated sTnC. We have demonstrated that it is possible to place a single lanthanide ion into site II of calcium-saturated cTnC, and to monitor the formation of the species by simple one-dimensional (1D) NMR. By using this method, we have been able to extract orientation parameters from ^{15}N-cTnI$_{128-147}$ bound to unlabeled cTnC containing a single bound Yb^{3+} ion in site II.

EXPERIMENTAL PROCEDURES

E.1 Sample Preparation

Recombinant human cTnC (residues 1–161) with the mutations C35S and C84S (denoted cTnC [C35S, C84S]), and chicken sTnC (residues 1–163), in which C101 was not mutated, were used in this study. Hereafter, the cysteine-less version of cTnC will be referred to as cTnC for simplicity. The cTnC (C35S, C84S) expression vector was constructed as described previously (*27*) and was expressed, purified, and isotopically labeled in BL21 (DE3) pLysS *E. coli* cells, as described previously for ^{15}N-cTnC N-lobe (*40*) with decalcification performed by using 200mM EDTA instead of 100mM. The higher concentration of EDTA helped to remove any tightly bound metal ions in the C-lobe of cTnC. The process used in the expression and purification of the chicken sTnC is as previously described (*41*), with the exception that iodoacetamide was not added to prevent Cys oxidation. The cTnI peptide, cTnI$_{128-147}$ **TQKIFDLRGKFKRPTLRRVR-hSer,** was prepared as a fusion protein with GB1 by using recombinant techniques and separated via CNBr cleavage. The expression and purification of the ^{15}N-cTnI$_{128-147}$ in BL21 DE3 (pLysS) cells were as described previously (*28, 42, 43*). The peptide sequence was confirmed by DNA sequencing, and the mass was verified by MALDI-TOF mass spectrometry.

All NMR samples were 500μL in volume. The buffer conditions were 100mM KCl, 10mM imidazole, 0.2mM 2,2-dimethyl-2-silapentanesulfonic acid (DSS), and 0.01% NaN_3 in 90% H_2O/10% D_2O, and the pH was 6.7. The concentration of the apo ^{15}N-cTnC sample used for Ce^{3+} titration was determined by amino acid analysis to be 0.87mM (*44*). By using the same methods, the concentration of the ^{15}N-cTnC sample used for the cTnI$_{128-147}$ titration was 0.56mM. Each of the samples contains ~3 to 5mM $CaCl_2$.

E.2 Lanthanide titrations of cTnC

Stock solutions of 100mM $CaCl_2$, standardized by atomic absorption spectroscopy, and 284mM $CeCl_3$ were used for the titration. The stock solution of $CeCl_3$ was prepared from solid by dissolution of $CeCl_3 \bullet 5H_2O$ in distilled H_2O. A 100-fold diluted sample of the stock solution, combined with 0.05% xylenol orange dye, was then calibrated via titration with EDTA. To an NMR tube containing a 500*u*L sample of 0.89mM ^{15}N-cTnC, 13.5*u*L of stock $CaCl_2$ solution diluted to 89mM were added and thoroughly mixed. For a total of 30 additions following, 0.5*u*L aliquots of stock $CeCl_3$ solution diluted to 0.89mM were added, mixing thoroughly after each addition. All protein concentrations were determined via amino acid analysis. The total volume increase was 28.5*u*L, and the change in protein concentration due to dilution was taken into account for data analysis. An acidic change in pH of ~0.3 units due to the Ca^{2+} and Ce^{3+} additions was noted. Both 1D 1H and 2D {^{1}H, ^{15}N}-HSQC NMR spectra were acquired at 500 MHz at every titration point. This was repeated for both $TbCl_3$ and $YbCl_3$, with concentrations of 158.08 and 497.13mM, respectively.

E.3 Lanthanide titration of sTnC

Stock solutions of 100mM $CaCl_2$ and 284mM $CeCl_3$ were used for the titration and were standardized as above. To an NMR tube containing a 500*u*L sample of 0.92mM ^{15}N-sTnC, 18*u*L of stock $CaCl_2$ solution diluted to 92mM was added and thoroughly mixed. For a total of 30 additions following, 0.5*u*L aliquots of stock $CeCl_3$ solution diluted to 0.92mM were added, with mixing thoroughly after each addition. All protein concentrations were determined by amino acid analysis. The total volume increase was 38*u*L, and the change in protein concentration due to dilution was taken into account for data analysis. An acidic change in pH of ~0.2 units due to the Ca^{2+} and Ce^{3+} additions was noted. Both 1D 1H and 2D {^{1}H, ^{15}N}-HSQC NMR spectra were acquired at 500 MHz at every titration point. This titration was repeated up to 15 additions with $TbCl_3$, from a stock solution of concentration 158.08mM standardized in the same manner as that of $CeCl_3$, using a protein concentration of 0.61mM and a $TbCl_3$ stock dilution to 61mM.

E.4 Yb^{3+} Titration of $cTnC{\bullet}3Ca^{2+}{\bullet}^{15}N\text{-}cTnI_{128\text{-}147}$

Stock solutions of 100mM $CaCl_2$ and 497mM $YbCl_3$, and 50mM ^{15}N-cTnI$_{128\text{-}147}$ were used for the titration. The concentration of the ^{15}N-cTnI$_{128\text{-}147}$ solution was determined by amino acid analysis. To an NMR tube containing a 500*u*L sample of 0.50 mM ^{15}N-cTnC, 13.5*u*L of stock CaCl2 solution diluted to 56 mM were added and thoroughly mixed. Following this, five 1*u*L aliquots of ^{15}N-cTnI$_{128\text{-}147}$ solution were added, mixing thoroughly after each addition, up to one molar equivalent of cTnC. Both 1D ^{1}H and 2D {^{1}H, ^{15}N}-HSQC NMR spectra were acquired at 500 MHz at every titration point. Stock YbCl3 solution was diluted to 60mM and added in 0.2*u*L additions while monitoring G70, G110, and G146 resonances by 1D 1H NMR at 500 MHz. The titration was stopped after 12 additions when the first sign of a loss of intensity of G110 was noted to minimize the amount of Yb^{3+} in the C-lobe. The total volume increase was 20.9*u*L, and the change in protein concentration due to dilution was taken into account for data analysis. An acidic change in pH of ~0.3 units due to the Ca^{2+}, cTnI$_{128\text{-}147}$, and Yb^{3+} additions was noted. Both 2D {^{1}H, ^{15}N}-HSQC and 2D {^{1}H, ^{15}N}-IPAP NMR spectra were acquired at 600 and 800 MHz.

E.5 NMR spectroscopy

All of the NMR spectra were obtained at 30°C by using Varian Unity [illegible]00 MHz and Varian INOVA 500 MHz and 800 MHz spectrometers. 2D {^{1}H, ^{15}N}-HSQC NMR spectra were acquired by using the sensitivity-enhanced gradient pulse scheme developed by Lewis Kay and coworkers (*45, 46*). The ^{1}H and ^{15}N sweep widths were 7500 and 1500 Hz on the 500 MHz spectrometer; 8000 and 1650 Hz on the 600 MHz spectrometer; and 12000 and 2200 Hz on the 800-MHz spectrometer. 2D {^{1}H, ^{15}N}-IPAP NMR spectra were acquired by using 3919 WATERGATE suppression pulse scheme (*47*). All spectra were processed and analyzed by using VNMR (Varian Associates), NMRPipe (*48*), and NMRView (*49*) and were referenced according to the IUPAC conventions.

RESULTS

The purpose of this study was to determine the potential for using paramagnetic lanthanide ions to partially align troponin C in solution as a tool for the structure determination of bound TnI peptides. To do so, it was necessary to determine the site specificity for lanthanide addition to troponin C as a prerequisite for structural applications.

E.1 Lanthanide titration of Ca^{2+}-saturated cTnC

In this study two-dimensional (2D) $\{^1H, ^{15}N\}$-HSQC NMR spectroscopy was used to monitor the binding order of different lanthanide ions to ^{15}N-cTnC. Previous studies have demonstrated the utility of 2D $\{^1H, ^{15}N\}$-HSQC NMR in characterizing Ca^{2+} and TnI peptide binding to both sTnC and cTnC (*50*). The 2D $\{^1H, ^{15}N\}$-HSQC spectrum of Ca^{2+}-saturated cTnC has previously been assigned (*21*) and was used to determine which residues were affected by pseudo-contact shifting and/or paramagnetic broadening during the lanthanide titrations. Figure E-1 depicts the same expanded region from the 2D $\{^1H, ^{15}N\}$-HSQC spectrum for each of the four species that were generated during the titrations of cTnC•3Ca^{2+} with $CeCl_3$, $TbCl_3$, and $YbCl_3$. Figure E-1A shows the holo-cTnC spectrum at the beginning of the titration, with the labeled cross-peaks corresponding to amino acid residues that occur in the three active binding loops of cTnC. Residues E66, V72, and D73 are located in site II; I112 and D113 in site III; and I148, D149, and F153 in site IV. Figure E-1(B-D) shows the same spectral region after one molar equivalent of Ce^{3+}, Tb^{3+}, and Yb^{3+}, respectively, have been titrated into holo-cTnC. In this figure asterisks mark those residues with resonances that had disappeared from their original positions and/or broadened beyond detectability due to close proximity to a bound lanthanide. In all cases, those resonances that had disappeared corresponded to the residues found in the site II binding loop, whereas those that corresponded to sites III and IV had not shifted noticeably. This was an indication that all three lanthanide ions had preferentially bound to, and displaced Ca^{2+} from, the N-lobe binding site. Note that some of the unassigned resonances in the upper right had also shifted or broadened upon addition of lanthanide. These resonances are also located in the N-lobe but are not part of the binding loop. Further note the change in intensity of the C-lobe binding loop residues in Figure E-1(C, D). Because the degree of shifting and broadening due to lanthanide binding is dependent upon $1/r^3$ and $1/r^6$, respectively (*51*), we would not expect this broadening to come from lanthanides bound to the N-lobe. Thus, the change in intensity demonstrated that there was partial binding of lanthanide to sites III and IV by the time one molar equivalent had been reached in the titration.

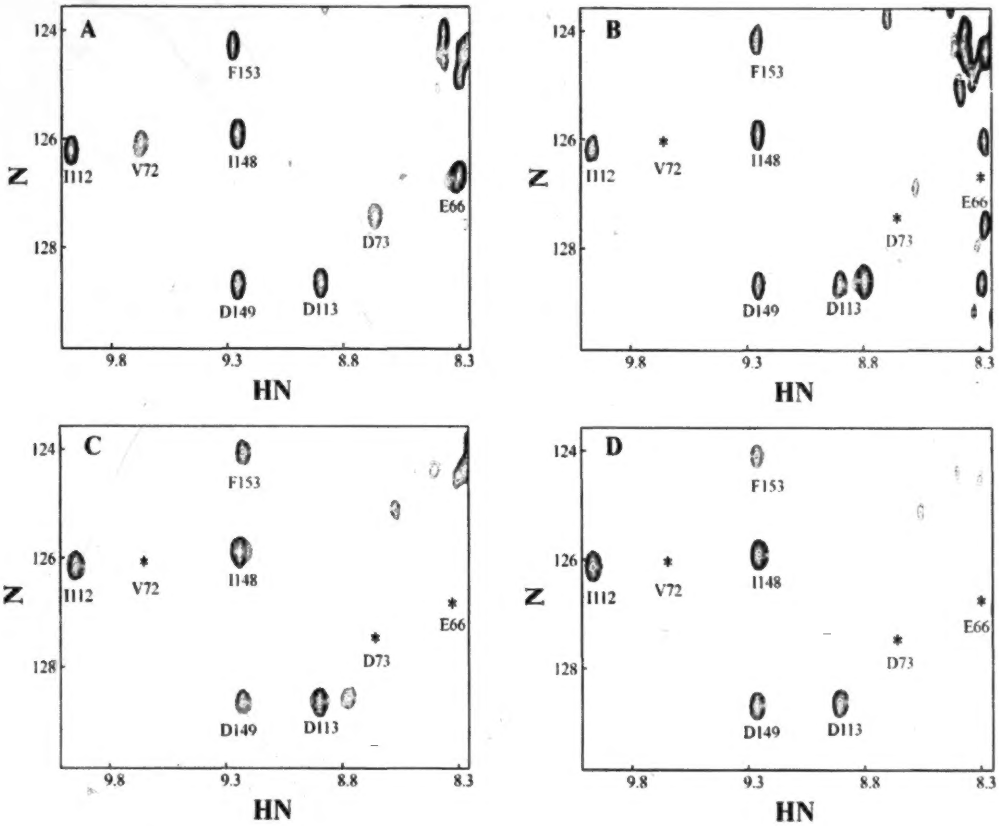


Figure E-1. **HSQC lanthanide titration of cTnC.** Portions of a 2D {^{1}H, ^{15}N}-HSQC spectra of ^{15}N-cTnC saturated with Ca^{2+} (*A*), subsequently titrated to one molar equivalent with Ce^{3+} (*B*), Tb^{3+} (*C*), or Yb^{3+} (*D*). Asterisks mark the position of N-lobe binding site II resonances that have either shifted or broadened beyond detection after the addition of one molar equivalent of lanthanide to holo-cTnC

Figure E-2 shows the results of the continued titration of Ca^{2+}-saturated cTnC with Ce^{3+} beyond the first molar equivalent. Here we observed the glycine residues located in the middle of the binding loops, G110 and G146, corresponding to sites III and IV respectively. Although Ca^{2+} is displaced from both sites, the site III resonances diminished faster than their site IV counterparts, and by the time two molar equivalents of Ce^{3+} had been added, there was no longer a detectable resonance from G110. This indicated a preference for site III over site IV.

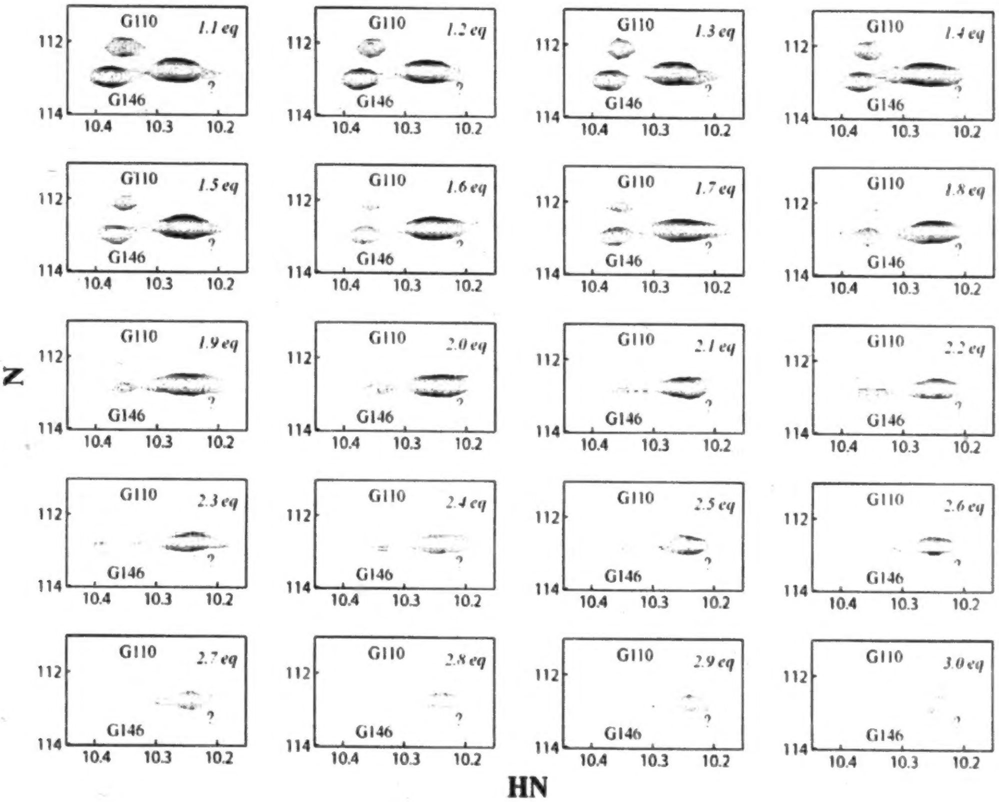


Figure E-2. **Ce^{3+} titration of cTnC.** Portions of a 2D {^1H, ^{15}N}-HSQC spectra of ^{15}N-cTnC•Ce^{3+}•$2Ca^{2+}$ depicting the downfield Gly residues of the C-lobe binding loops III and IV. This series demonstrates the binding order of Ce^{3+} to the C-lobe sites after the binding of one molar equivalent of Ce^{3+} to holo-cTnC. The titration progresses from the *upper left* box at 1.1 eq and proceeds to the *lower right* by 0.1 eq steps. The two C-lobe Gly residues are labeled, and an additional unassigned residue is labeled with a question mark. This residue moved into this region upon the binding of the first equivalent of Ce^{3+} and is likely an N-lobe resonance.

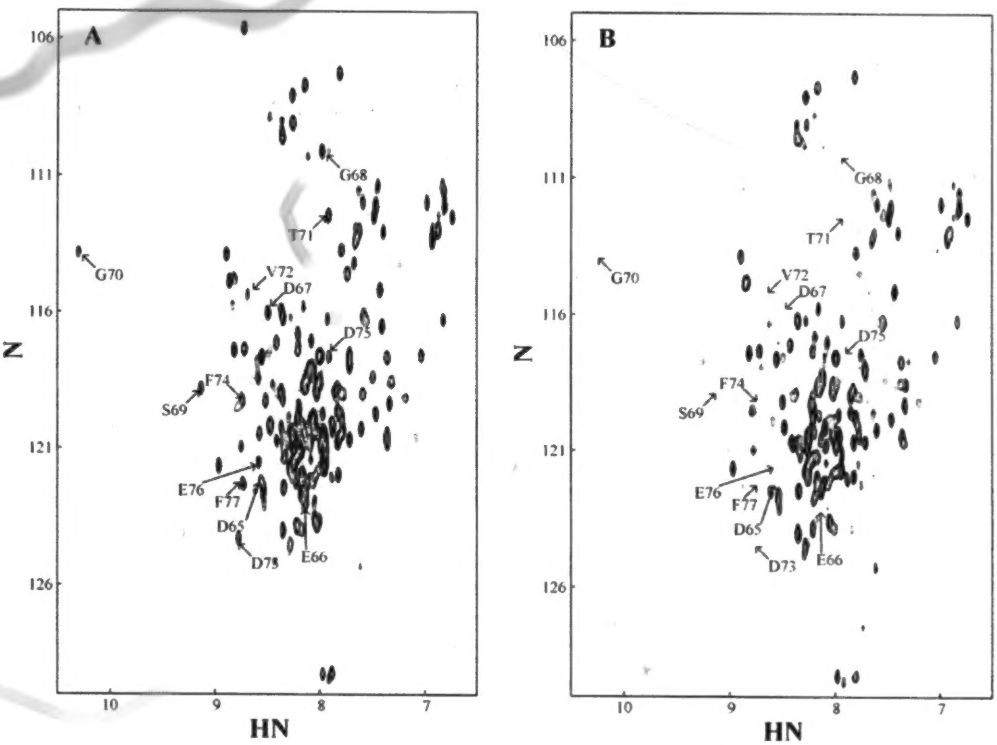


Figure E-3. **cTnC HSQC.** 2D {^{1}H, ^{15}N}-HSQC spectrum of both ^{15}N-apo-cTnC (*A*) and ^{15}N-cTnC bound to one molar equivalent of Ce^{3+} (*B*). Arrows point to the apo-positions of the site II binding loop resonances in both spectra.

In the titrations of holo-cTnC with lanthanides, a displacement of the weakest bound 2+ cation first by a stronger 3+ cation appeared to be occurring. To determine which site was occupied first in the absence of competing Ca^{2+}, apo-^{15}N-cTnC was titrated with CeCl$_3$. Figure E-3 depicts the 2D {^{1}H, ^{15}N}-HSQC spectrum of both apo-cTnC (Figure E-3A) and cTnC bound to one molar equivalent of Ce^{3+} (Figure E-3B). In each case, the resonances corresponding to the residues of the N-lobe binding site are indicated with arrows. Note that after the addition of 1 eq of Ce^{3+}, these resonances had either shifted from their initial positions or disappeared, implying that binding had occurred in site II. This is significant in that it demonstrated a preference for the N-lobe binding site over the C-lobe sites, even when there were no competing ions to displace.

Also note that that in apo-cTnC the C-lobe is unstructured and thus, resonances for the C-lobe Gly residues 110 and 146 were not observed.

E.3 Lanthanide titration of Ca2+-saturated sTnC

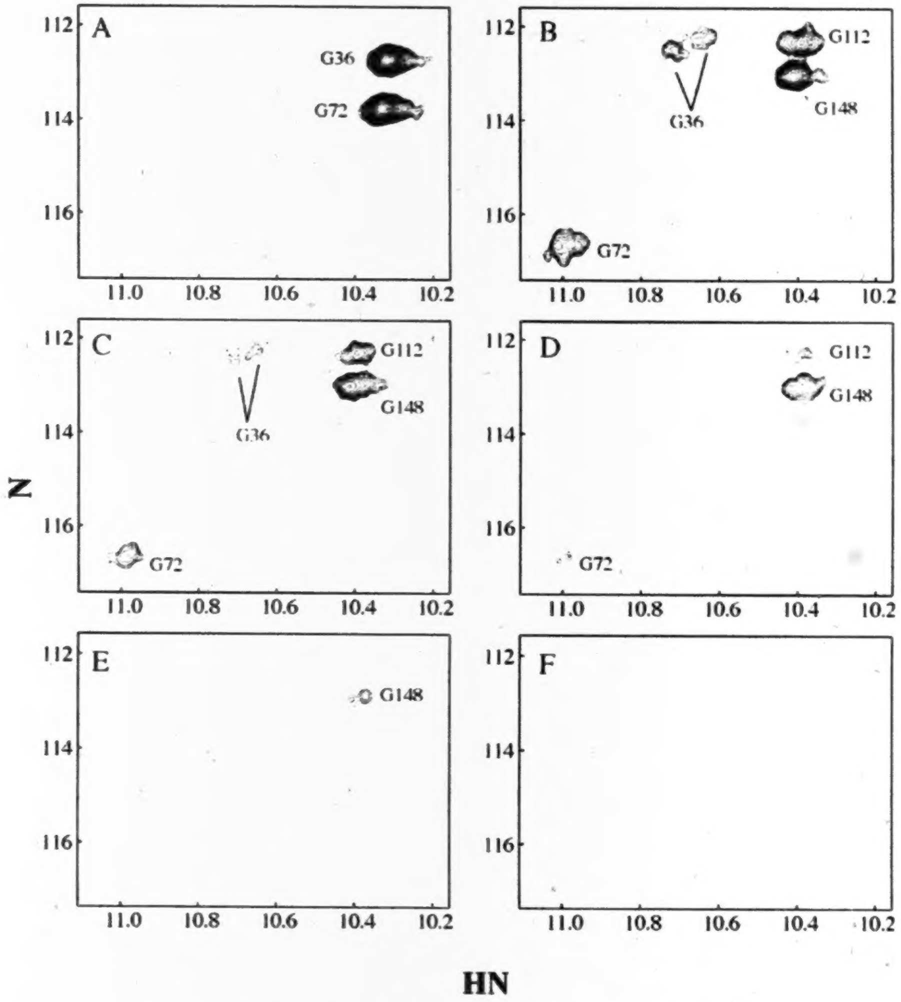


Figure E-4. **Ce^{3+} titration of sTnC.** Portions of a 2D {^{1}H, ^{15}N}-HSQC spectrum showing several points in the titration of 0.92mM 15N-sTnC with Ce^{3+}. The area containing downfield Gly residues of the ion binding loops is expanded, and the panels follow through the titration from apo-saturated (*A*) to Ca^{2+}-saturated (*B*) through one, two, three, and four molar equivalents (*C–F*, respectively). All visible residues are labeled. Note that the splitting in the N-lobe Gly resonances is likely due to Cys cross-bridge formation.

In the results above, a demonstration of an N-lobe binding preference by lanthanides for cTnC can be seen. In previous fluorescence studies with sTnC, however, the binding preference was found to be the same among lanthanides as for Ca^{2+} (*32*). In an attempt to understand the differences between the cardiac and skeletal isoforms with respect to their lanthanide binding order preferences, a titration of sTnC was performed in a manner identical to the cTnC titrations. Figure E-4 shows an expanded region of the 2D $\{^1H, ^{15}N\}$-HSQC spectrum for several points in the titration of sTnC with Ce^{3+}. This region shows the downfield-shifted Gly residues found in the center of each binding loop in sTnC, and assignments were made based on previously assigned 2D $\{^1H, ^{15}N\}$-HSQC spectra (*52*). Figures E-4(A, B) show the apo- and Ca^{2+}-saturated states and Figures E-4(C-F) show the changes to the spectrum after the addition of one, two, three, or four molar equivalents of Ce^{3+}, respectively. As Ce^{3+} was added to the solution, the N-lobe peaks were the first to disappear, followed by the C-lobe. Note also that the N-lobe peaks were split, G36 more so than G72. This cause of the splitting is likely due to Cys cross-bridge formation at high concentrations dimerizing the protein at the N-lobe because neither DTT nor β-mercaptoethanol was added to the protein solution (*53*). Again, as in the cTnC spectra, the C-lobe is unfolded in the apo-state; thus, the C-lobe Gly resonances were not observed in Figure E-4A.

E.4 Orientation of cTnC•Yb^{3+}•$2Ca^{2+}$ and RDC measurements of bound ^{15}N-$cTnI_{128\text{-}147}$

The 1D ^1H NMR spectra shown in Figure E-5(A, B) demonstrate the selective binding of the Yb^{3+} ion to site II of calcium-saturated cTnC. Note that the degree that the C-lobe cTnC Gly resonances were diminished was kept to a minimum by halting the addition of Yb^{3+} to the cTnC•$3Ca^{2+}$•^{15}N-$cTnI_{127\text{-}148}$ at the first sign of a reduction in intensity. This was done to ensure that the least amount of Yb^{3+} possible would be present in the C-lobe, so that the resonances from the bound peptide would not be broadened. To that end, only ~0.6 molar equivalents of Yb^{3+} were added to the TnC–TnI complex. A $\{^1H, ^{15}N\}$-IPAP spectra from an 800 MHz spectrometer is shown in Figure E-5C. In it, the resonances of ^{15}N-$cTnI_{127\text{-}148}$ residues (assigned as per (*42*)) are shown with their IPAP peaks connected with a dashed line, and the measured RDC (in hertz) noted. What is actually observed in the spectrum is a single averaged peak for each peptide residue, due to the fast exchange rate of the bound cTnI peptide between the Yb^{3+}•$2Ca^{2+}$ and $3Ca^{2+}$ species of cTnC. To extract the RDCs from the measured couplings, measurements were taken from both an unaligned and a partially aligned sample and subtracted to yield the residual coupling. This is done because the measured coupling is actually a sum of the $^1J_{NH}$ scalar coupling and the RDC (*54*). The RDC itself is dependent upon the square of the applied magnetic field (B_o) as shown by

$$\begin{aligned} RDC(\theta,\phi) \;&=\; -(B_o^2/15kT) \\ &[\gamma_H\gamma_N h/(4\pi^2 r_{HN}^3)] \times [\Delta\chi_a(3\ cos^2\theta - 1) \\ &+ 3/2\Delta\chi_r(sin^2\ \theta cos2\phi)] \end{aligned}$$

where *a* and *r* are the axial and rhombic components of the magnetic susceptibility tensor (χ), and are the angular coordinates of the orientation of the N-H bond vector within the principal axis system of χ (*54*).

When the absolute value of the measured RDCs are plotted against the amino acid sequence of $cTnI_{127-148}$ (Figure E-6), the strongest RDCs are found near the center of the peptide, whereas the weakest are found near the ends. In particular, this center area with the strongest RDCs has been found to be the only structured part of the peptide when bound to cTnC (*42*). This trend is to be expected because relatively unstructured regions of a polypeptide would experience much less average alignment than do structured regions. When the results of similar experiments on a 600 MHz spectrometer (data not shown) were compared with those above, the differences in measured RDCs were strictly dependent on the expected difference due to a weaker B_o. This demonstrated the large errors that are present in taking RDC measurements as the extraction of the data depends largely upon the accuracy of the measure of peak centers. Because of these errors and the relatively small RDCs observed, a structure was not calculated from the RDC data presented in Figure E-6.

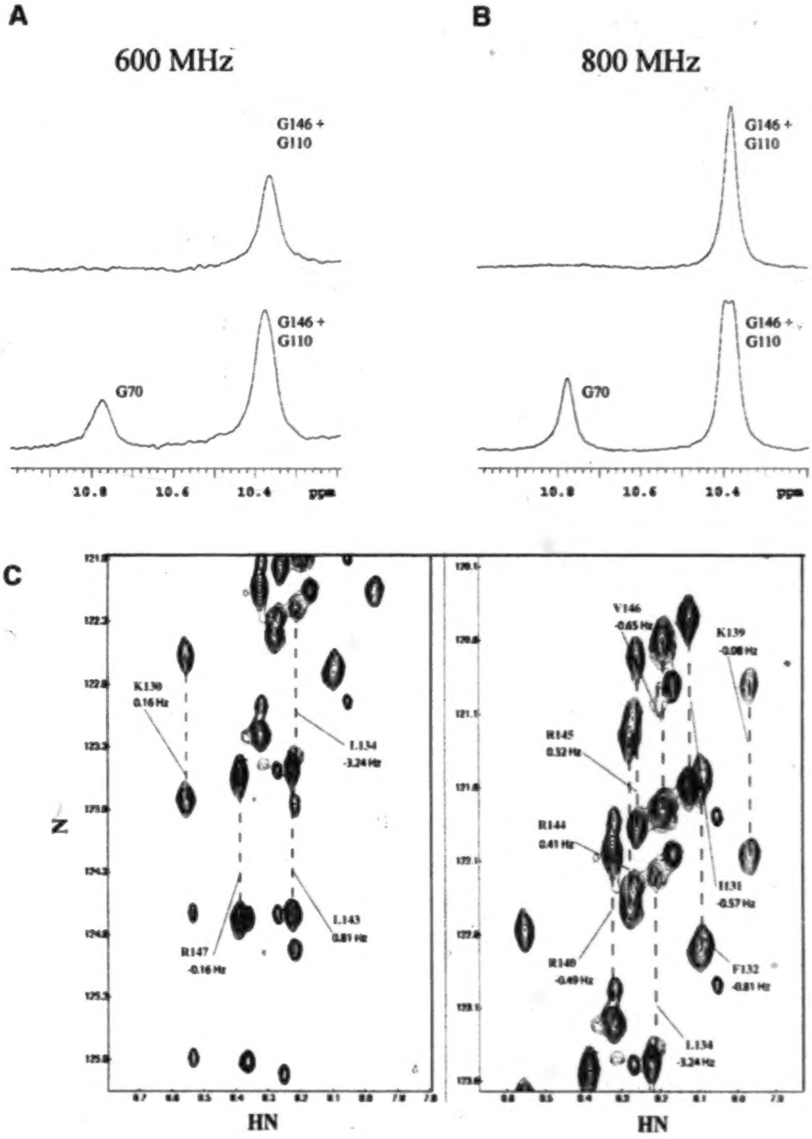


Figure E-5. **RDC measurements of bound inhibitory region.** Expanded region of 1D NMR spectra at 600 MHz (*A*) and 800 MHz (*B*) of unlabeled cTnC, as well as a 2D {^{1}H, ^{15}N}-IPAP spectrum also at 800 MHz (*C*) of ^{15}N-cTnI$_{127\text{-}148}$ bound to unlabeled cTnC. (*A, B*) The successful addition of Yb^{3+} solely to the N-lobe site II, with the *lower* spectrum in each depicting the downfield Gly residues of cTnC•3Ca^{2+}, and the *upper* spectrum depicting the same region of cTnC•Yb^{3+}•2Ca^{2+}. (*C*) The splitting of the ^{15}N-cTnI$_{127\text{-}148}$ resonances and the measured RDCs for each residue. For each pair of resonances connected by a dashed line, the *upper* resonance is the inphase resonance and the *lower* resonance is antiphase.

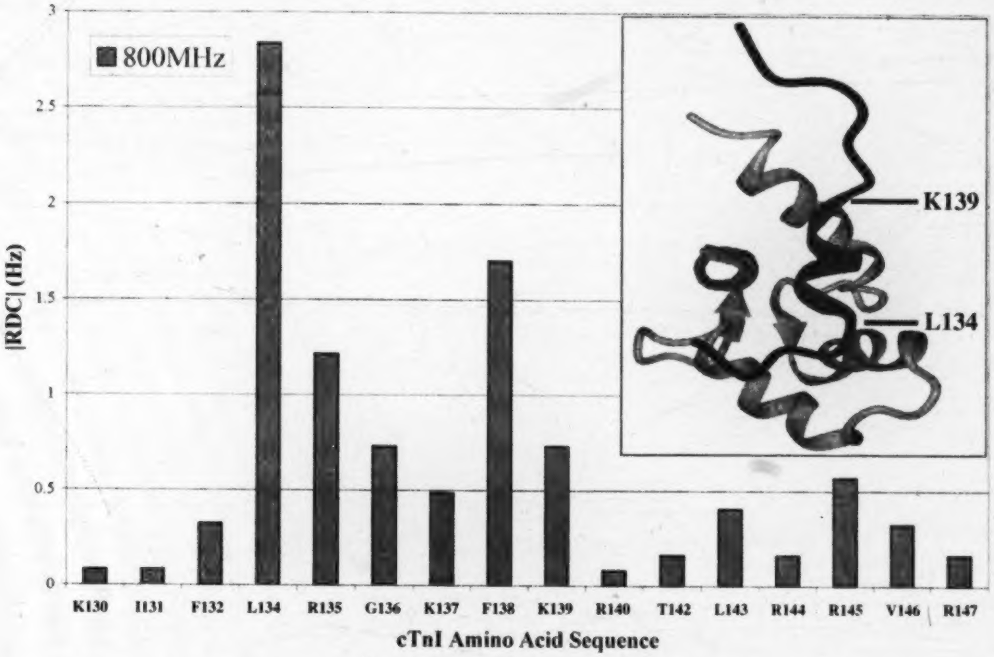


Figure E-6. **RDC plot of bound inhibitory region.** ^{15}N-cTnI$_{127\text{-}148}$ at 800 MHz versus the amino acid sequence of ^{15}N-cTnI$_{127\text{-}148}$. The *inset* structure of bound ^{15}N-cTnI$_{127\text{-}148}$ to the C-lobe of cTnC is taken from (*42*).

DISCUSSION

In this study the binding of the lanthanide ions tested to cTnC has been shown to be in the order of site II → site III → site IV. This is in direct contrast to the observed binding order of Ca^{2+}, which has been demonstrated to be III + IV → II (*16*). The same can be said for the skeletal isoform, as the binding order of Ca^{2+} there has been demonstrated to be III + IV → I + II (*55*), yet we observe the opposite for Ce^{3+}. This latter observation is not in agreement with previous fluorescence measurements that determined the Tb^{3+} binding order to be the same as that of Ca^{2+} (*32*). Given that two of the ions tested, Ce^{3+} and Yb^{3+}, span the lanthanide series, it is assumed that this behavior is the same for all lanthanides in that series. To understand this phenomenon, the crystal structures of calcium saturated chicken cTnC and rabbit sTnC were examined. Although there are differences in the primary sequence between the proteins used in these experiments and those of the crystal structures, the differences have no effect on the electrostatics of the binding loops.

The 12 residues that compose each active binding loop in both cTnC and sTnC are shown in *Table E-1.* The data in these tables were generated by first calculating the net charge of each ion binding loop in both cTnC and sTnC at physiological pH. This value is listed beneath each residue and the total summed at the end of the line. The crystal structures of both chicken cTnC (*Table E-1A* and rabbit sTnC *Table E-1B*) were examined, and distances from the bound ion to the various charged loop residues were measured. By dividing the residue charge in Coulombs over the distance measured, a charge/distance ratio was obtained. This value is directly proportional to the electrostatic binding energy on each bound ion

$$E_{total} = \frac{q_M}{4\pi\varepsilon_o}\left(\sum_{i=1}^{12}\frac{q_i}{r_{iM}}\right)$$

where E_{total} is the electrostatic energy from the loop on a particular bound ion, ε_o is the permittivity of free space, q_M is the charge on the ion, q_i is the charge on loop residue i, and r_{iM} is the distance between the bound ion and the charge center of each residue. In the case of Glu and Asp, the charge center was placed on the side-chain carboxyl carbon; for Arg, the charge was placed on the average distance as measured from the two terminal nitrogens of the side chain; and for Lys, the charge was placed on the terminal nitrogen of the side chain. Each value is listed below the residue it corresponds to, and the total is summed at the end of the row. It should be noted that this electrostatic model does not take into consideration dielectric compensation of the protein and solvent, and it assumes that all groups are fully ionized and the structure does not vary.

Table E-I

The active binding loops of (A) chicken cardiac and (B) rabbit skeletal troponin C

A.							cTnC						
Site II	D	E	D	G	S	G	T	V	D	F	D	E	F	Total
Residues (62–74)													
Charge	-1	-1	-1	0	0	0	0	0	-1	0	-1	-1	0	-6
$q_i/r_{iM} \times 10^{-9}$ (C/m)	-0.48	-0.19	-0.48	—	—	—	—	—	-0.33	—	-0.17	-0.56	—	-2.2
Site III	D	K	N	A	D	G	Y	I	D	L	E	E	L	Total
Residues (105–117)													
Charge	-1	+1	0	0	-1	0	0	0	-1	0	-1	-1	0	-4
$q_i/r_{iM} \times 10^{-9}$ (C/m)	-0.48	+0.16	—	—	-0.49	—	—	—	-0.31	—	-0.15	-0.54	—	-1.8
Site IV	D	K	N	N	D	G	R	I	D	Y	D	E	F	Total
Residues (141–153)													
Charge	-1	+1	0	0	-1	0	+1	0	-1	0	-1	-1	0	-3
$q_i/r_{iM} \times 10^{-9}$ (C/m)	-0.45	+0.17	—	—	-0.50	—	+0.22	—	-0.32	—	-0.16	-0.55	—	-1.6
B.							sTnC							
Site I	D	A	D	G	G	G	D	I	S	V	K	E	V	Total
Residues (28–40)													
Charge	-1	0	-1	0	0	0	-1	0	0	0	+1	-1	0	-3
$q_i/r_{iM} \times 10^{-9}$ (C/m)	-0.46	—	-0.48	—	—	—	-0.30	—	—	—	+0.34	-0.56	—	-1.5
Site II	D	E	D	G	S	G	T	I	D	F	E	E	F	Total
Residues (64–76)													
Charge	-1	-1	-1	0	0	0	0	0	-1	0	-1	-1	0	-6
$q_i/r_{iM} \times 10^{-9}$ (C/m)	-0.45	-0.18	-0.47	—	—	—	—	—	-0.32	—	-0.16	-0.54	—	-2.1
Site III	D	R	N	A	D	G	Y	I	D	A	E	E	L	Total
Residues (104–116)													
Charge	-1	+1	0	0	-1	0	0	0	-1	0	-1	-1	0	-4
$q_i/r_{iM} \times 10^{-9}$ (C/m)	-0.47	+0.16	—	—	-0.48	—	—	—	-0.32	—	-0.16	-0.57	—	-1.8
Site IV	D	K	N	N	D	G	R	I	D	F	D	E	F	Total
Residues (140–152)													
Charge	-1	+1	0	0	-1	0	+1	0	-1	0	-1	-1	0	-3
$q_i/r_{iM} \times 10^{-9}$ (C/m)	-0.43	+0.17	—	—	-0.49	—	+0.22	—	-0.32	—	-0.17	-0.58	—	-1.6

For each residue, the charge and charge/distance ratio is listed and then totaled for each loop at the end of the line. The charge of each residue is given at physiological pH. The charge/distance ratio is calculated by using the formula *qi/riM*, where *qi* is the charge of residue *i* in Coulombs, and *riM* is the distance from residue *i* to the bound metal ion as measured from the carbonyl carbon of Asp and Glu or the terminal nitrogen of Arg and Lys. For Arg, the distance is the average of the measure from both terminal nitrogens.

When the loop is examined, it can be seen that the phenomenon of lanthanide binding order follows the decrease of net negative charge and charge/distance ratio of the residues. This is also true when considering the effective charge at the ion binding site. This seems to be an adequate explanation for the observed results, given that the similar size of the lanthanide ions to the calcium ion would indicate that this preference for the lanthanide by cTnC is driven solely by the greater positive charge density of the lanthanide. This has been proposed to be the case in other systems (*56*). Thus, it seems quite reasonable to expect such a positive charge density–driven replacement of a bound ion to result in sites with a greater net negative charge being filled first.

The results of the Ce^{3+} titration of sTnC, however, seem to differ with this explanation. In *Table E-1B*, the net charge of the ion-binding loops of sTnC seems to be in the order of II → III → I and IV, from most to least negative. When the effective charge at the bound ion is examined, the trend appears to be the same. Given the above charge/density explanation, one would expect to see G72 disappearing first, followed by G112, but in Figure E-4 we see G36 vanishing followed by G72. Note that G72 also begins to vanish before G36 is completely gone, as does G112, indicating that although a preference is present, it is quite small. This result contradicts the previously published fluorescence-based work on sTnC, which states that the binding order would be C-lobe first then N-lobe (*32*). The aforementioned work, however, could not distinguish between Y10 and Y109 fluorescence in rabbit sTnC, which was later shown to have disparate contributions to the total observed fluorescence (*57*). It would appear then that although electrostatics plays a significant role in the binding of lanthanides to troponin C, there are other factors, such as binding cooperativity, ligand coordination, or even charge repulsion from bound ions, that influence the order in which these ions bind to the protein.

With one bound Yb^{3+}, it was possible to confer partial alignment on bound ^{15}N-$cTnI_{127\text{-}148}$ and measure small RDCs. The size of the RDCs could be due to several factors, given the conditions under which they were measured. First, the amount of Yb^{3+} bound to cTnC was 0.6 molar equivalents, and so, only ~60% of the molecules in solution would experience some sort of partial alignment. One could choose a paramagnetic ion with a higher magnetic moment such as Ho^{3+} or Dy^{3+}, which could induce a higher degree of alignment upon binding to the protein. Another factor could be that $cTnI_{127\text{-}148}$ will possess only one structured area when bound to cTnC. The remaining unstructured area, the C-terminal end of the peptide, has a great deal of mobility, and thus, even if an overall magnetic susceptibility tensor is conferred on the complex, the degree of local alignment of these residues in the peptide could be minimal. Finally, the degree of flexibility in the linker helix of cTnC could have an effect on the degree of local alignment of the peptide residues. This could be overcome by using a longer peptide or intact cTnI to add rigidity to the complex by tying the N- and C-lobes of cTnC. Despite the lack of RDCs large enough to calculate a three-dimensional structure, it was possible to note the areas of the peptide that were more structured. (Figure E-6) An outline of the secondary structure of the peptide can be seen even from the weak RDCs collected when plotted against the sequence of the peptide.

CONCLUSIONS

In the end, by determining the order of lanthanide binding, it is possible to create a troponin species containing one paramagnetic center, which can be used to confer partial alignment on a bound ligand to obtain orientation-based restraints for structure calculation. Although it was relatively easy to determine the binding order necessary for the creation of the paramagnetic species, predicting this binding order is not quite as simple. Although understanding the electrostatics of the ion-binding loop is essential, there are many other factors that also have a hand in governing the binding of an ion that were not addressed in the proposed model. The degree to which these factors influence this phenomenon determines how easily simple electrostatics can be used to predict ion binding order in any such protein. Also, from what was learned about lanthanide occupancy in the troponin isoforms, it can be concluded that the behavior of the isoforms should not be assumed to be similar with respect to their ion-binding abilities. This also speaks to the manner by which different ions bind to each isoform. When metal ions are used so often in biophysical chemistry as probes, or other means to obtain measurements in proteins, care should be taken in understanding their binding order sufficiently.

REFERENCES

1. Perry, S. V. (1999) *Mol Cell Biochem 190*, 9-32.
2. Perry, S. V. (2003) *J Muscle Res Cell Motil 24*, 593-6.
3. Tobacman, L. S. (1996) *Annu. Rev. Physiol. 58*, 447-481.
4. Farah, C. S., and Reinach, F. C. (1995) *FASEB J. 9*, 755-767.
5. Squire, J. M., and Morris, E. P. (1998) *FASEB Journal 12*, 761-771.
6. Filatov, V. L., Katrukha, A. G., Bulargina, T. V., and Gusev, N. B. (1999) *Biochemistry (Mosc). 64*, 969-85.
7. Clark, K. A., McElhinny, A. S., Beckerle, M. C., and Gregorio, C. C. (2002) *Annu Rev Cell Dev Biol 18*, 637-706.
8. Gordon, A. M., Homsher, E., and Regnier, M. (2000) *Physiological Reviews 80*, 853-924.
9. Solaro, R. J. (2003) *Adv Exp Med Biol 538*, 389-401; discussion 401-2.
10. Solaro, R. J. (1999) *Circ Res 84*, 122-4.
11. Kobayashi, T., and Solaro, R. J. (2004) *Annu Rev Physiol.*
12. Sumandea, M. P., Burkart, E. M., Kobayashi, T., De Tombe, P. P., and Solaro, R. J. (2004) *Ann N Y Acad Sci 1015*, 39-52.
13. Perry, S. V. (1999) *Mol. Cell. Biochem. 190*, 9-32.
14. Geeves, M. A., and Holmes, K. C. (1999) *Annu Rev Biochem 68*, 687-728.
15. Gagne, S. M., Li, M. X., McKay, R. T., and Sykes, B. D. (1998) *Biochem Cell Biol 76*, 302-12.
16. van Eerd, J.-P., and Takahashi, K. (1975) *Biochem. Biophys. Res. Commun. 64*, 122-127.
17. Robertson, S. P., Johnson, J. D., and Potter, J. D. (1981) *Biophys J 34*, 559-69.
18. Putkey, J. A., Sweeney, H. L., and Campbell, S. T. (1989) *J. Biol. Chem. 264*, 12370-8.
19. Sheng, Z., Pan, B. S., Miller, T. E., and Potter, J. D. (1992) *J. Biol. Chem. 267*, 25407-25413.
20. Gagné, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) *Nature Struct. Biol. 2*, 784-789.
21. Sia, S. K., Li, M. X., Spyracopoulos, L., Gagné, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) *J. Biol. Chem. 272*, 18216-18221.
22. Spyracopoulos, L., Li, M. X., Sia, S. K., Gagné, S. M., Chandra, M., Solaro, R. J., and Sykes, B. D. (1997) *Biochemistry 36*, 12138-12146.
23. Li, M. X., Spyracopoulos, L., and Sykes, B. D. (1999) *Biochemistry 38*, 8289-8298.
24. Spyracopoulos, L., McKay, R. T., and Sykes, B. D. (2000) *Biophys. J. 78*, 434A.
25. Solaro, R. J., and Rarick, H. M. (1998) *Circ Res 83*, 471-80.
26. McKay, R. T., Tripet, B. P., Pearlstone, J. R., Smillie, L. B., and Sykes, B. D. (1999) *Biochemistry 38*, 5478-89.
27. Li, M. X., Saude, E. J., Wang, X., Pearlstone, J. R., Smillie, L. B., and Sykes, B. D. (2002) *Eur Biophys J 31*, 245-56.
28. Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) *Biochemistry 41*, 7267-74.

29. Permi, P., Rosevear, P. R., and Annila, A. (2000) *J Biomol NMR 17*, 43-54.
30. Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1995) *Proc Natl Acad Sci U S A 92*, 9279-83.
31. Lee, L., and Sykes, B. D. (1983) *Biochemistry 22*, 4366-73.
32. Leavis, P. C., Nagy, B., Lehrer, S. S., Bialkowska, H., and Gergely, J. (1980) *Arch Biochem Biophys 200*, 17-21.
33. Wang, C. L., Leavis, P. C., Horrocks, W. D., Jr., and Gergely, J. (1981) *Biochemistry 20*, 2439-44.
34. Biekofsky, R. R., Muskett, F. W., Schmidt, J. M., Martin, S. R., Browne, J. P., Bayley, P. M., and Feeney, J. (1999) *FEBS Lett 460*, 519-26.
35. Bertini, I., Felli, I. C., and Luchinat, C. (2000) *Solution NMR of paramagnetic molecules*, Elsevier Science, The Netherlands.
36. Feeny, J., Birdsall, B., Bradbury, A. F., Biekofsky, R. R., and Bayley, P. M. (2001) *J Biomol NMR 21*, 41-8.
37. Barbieri, R., Bertini, I., Cavallaro, G., Lee, Y. M., Luchinat, C., and Rosato, A. (2002) *J Am Chem Soc 124*, 5581-7.
38. Dvoretsky, A., Abusamhadneh, E. M., Howarth, J. W., and Rosevear, P. R. (2002) *J Biol Chem 277*, 38565-70.
39. Johnson, J. D., Collins, J. H., Robertson, S. P., and Potter, J. D. (1980) *J. Biol. Chem. 255*, 9635-9640.
40. Li, M. X., Gagné, S. M., Spyracopoulos, L., Kloks, C. P. A. M., Audette, G., Chandra, M., Solaro, R. J., Smillie, L. B., and Sykes, B. D. (1997) *Biochemistry 36*, 12519-12525.
41. Slupsky, C. M., and Sykes, B. D. (1995) *Biochemistry 34*, 15953-64.
42. Lindhout, D. A., and Sykes, B. D. (2003) *J Biol Chem 278*, 27024-34.
43. Lindhout, D. A., Thiessen, A., Schieve, D., and Sykes, B. D. (2003) *Protein Sci 12*, 1786-91.
44. Smillie, L. B., and Nattriss, M. (1990) *PLC of peptides and proteins: Separation, analysis, and conformation. In Amino acid analyses of proteins and peptides: An overview*, CRC Press, Boca Raton, FL pg.809–821.
45. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.
46. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.
47. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson. 131*, 373-378.
48. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
49. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603-614.
50. McKay, R. T., Saltibus, L. F., Li, M. X., and Sykes, B. D. (2000) *Biochemistry 39*, 12731-12738.
51. Bertini, I., Felli, I. C., and Luchinat, C. (2000) *J Biomol NMR 18*, 347-55.
52. Slupsky, C. M., and Sykes, B. D. (1995) *Biochemistry 34*, 15953-15964.
53. Tsuda, S., Hasegawa, Y., Yoshida, M., Yagi, K., and Hikichi, K. (1988) *Biochemistry 27*, 4120-6.
54. Tjandra, N., Grzesiek, S., and Bax, A. (1996) *J Am Chem Soc 118*, 6264-6272.
55. Potter, J. D., and Gergely, J. (1975) *J. Biol. Chem. 250*, 4628-4633.

56. Atreya, H. S., Mukherjee, S., Chary, K. V., Lee, Y. M., and Luchinat, C. (2003) *Protein Sci 12*, 412-25.

57. Keleti, D., Rao, V. G., Su, H., Akella, A. B., Ding, X. L., and Gulati, J. (1994) *FEBS Lett 354*, 135-9.

APPENDIX F

NMR solution structure of a highly stable de novo heterodimeric coiled-coil

This appendix contains information regarding a synthetic heterodimeric coiled-coil, which is an NMR structural continuation of a project begun by Jennifer Litowski during her Ph. D. graduate career in the laboratory of Dr. Robert Hodges in the Department of Biochemistry at the University of Alberta. During her Ph. D., Jennifer studied and designed new coiled-coil motifs via stabilizing electrostatic interactions and hydrophobic contacts to address uses including affinity purification and biosensor applications. The majority of this work was previously described in the October 2002 issue of the *Journal of Biological Chemistry* [Designing heterodimeric two-stranded alpha-helical coiled-coils. Effects of hydrophobicity and alpha-helical propensity on protein folding, stability, and specificity - *J Biol Chem.* 2002 Oct 4;**277(40)**:37272-9]. NMR structural data and statistical analyses were performed by myself and Pascal Mercier in the Sykes laboratory, in the Department of Biochemistry at the University of Alberta. This work was published in the December 2004 issue of *Biopolymers* [*Biopolymers*. 2004 Dec 5;**75(5)**:367-75]. All authors involved in this body of work are Darrin A. Lindhout, Jennifer R. Litowski, Pascal Mercier, Robert S. Hodges and Brian D. Sykes. ***The data presented in this appendix is intended to supplement the discussion of the 35-residue coiled-coil motif found in the troponin complex between TnT and TnI on the IT arm. Specifics of coiled-coil interactions and affinities will be explored and contrasted, along with electrostatic interactions that are responsible for affinities and specificities.***

OVERVIEW

The NMR solution structure of a highly stable coiled-coil IAAL-E3/K3 has been solved. The E3/K3 coiled-coil is a 42-residue *de novo* designed coiled-coil comprising three heptad repeats per subunit, stabilized by hydrophobic contacts within the core and electrostatic interactions at the interface crossing the hydrophobic core which direct heterodimer formation. This E3/K3 domain has previously been shown to have high α-helical content as well as possessing a low dissociation constant (70*n*M). The E3/K3 structure is completely α-helical and is an archetypical coiled-coil in solution, as determined using a combination of ^{1}H-NOE and homology based structural restraints. This structure provides a structural framework for visualizing the important interactions for stability and specificity, which are key to protein engineering applications such as affinity purification and *de novo* design.

ABBREVIATIONS

Nuclear Magnetic Resonance - NMR, 2,2-Dimethyl-2-Silapentane-5-Sulfonate – DSS, sweep width – sw, number of scans – nt, number of increments – ni, Heteronuclear Single Quantum Coherence – HSQC, Total Correlation Spectroscopy – TOCSY, Nuclear Overhauser Effect Spectroscopy – NOESY, ^{2}H-trifluoroethanol – TFE-d_3, N^α-acetyl group – Ac, C^α-amide group – NH_2, Guanidine Hydrochloride – GdnHCl, Circular Dichroism – CD, Intensity of NOE crosspeak – I_{NOE}, Root Mean Square Deviation – RMSD

INTRODUCTION

The coiled-coil oligomerization domain has been known to occur in a wide variety of proteins, including muscle and motor proteins, DNA binding transcription factors, viral membrane fusion proteins and chaperone proteins (*1-11*). Coiled-coil domains contain two α-helical subunits wrapped around each other in a left handed twist that are stabilized by an extensive network of hydrophobic interactions within and electrostatic interactions across, the dimerization interface. This common structural motif for domain interactions has been widely exploited in evolutionary terms (*12*). The quaternary interactions present within coiled-coil proteins render them an excellent choice for structural studies by synthetic peptide approaches due to the relatively low molecular weights involved (<10kDa) and high-affinity association constants. The pursuit of structural studies via NMR spectroscopy is advantageous as coiled-coils are stable in solution over a wide-range of pH values, and are highly soluble due to the large number of charged surface residues. However the primary sequence composition is regular and unvaried, as well aggregation may occur at the concentrations required (≈1*m*M) for NMR solution structural determination. Spectral acquisition at high frequency (800MHz) is therefore required to maximize the resolution of the coiled-coil in solution, allowing for a higher number of ^{1}H-NMR assignments necessary for structural elucidation.

The structural features of the coiled-coil motif have been known for many years and have been extensively studied and reviewed (*5, 13, 14*). Coiled-coil complexes contain right-handed α-helical subunits interacting to form larger left-handed super structures (*15*). The primary amino acid sequences within the coiled-coil subunits contain a common repeat of 7 amino acids, which is denoted as **abcdefg**. The subunits can run parallel or anti-parallel to each other with the residues at positions **a** and **d** interacting to form the hydrophobic core. The residues present in position **a** generally favor a β-branched amino acid such as Ile or Val, and those present in position **d** generally favor a Leu due to packing considerations (*16-21*). Positions **e** and **g** within the heptad repeat contain charged residues such as Glu, Asp, Lys and/or Arg, which generally are placed as to be complementary to one another, allowing for beneficial electrostatic interactions between subunits in a *i* to *i*'+5 manner (*22, 23*). Positions **b**, **c**, and **f** are surface exposed, generally polar residues which aid to enhance protein solubility. These positions have been shown to be variable within protein sequences and may be important in further coiled-coil interactions with additional domains. They may also play a further role in coiled-coil stabilizing forces such as intra-chain electrostatics and α-helical propensity effects (*24*). The coiled-coil motif possesses a rod-like quaternary structure, and can exist as either homo and heterodimeric domains, as well as larger multimeric complexes containing 3, 4 and/or 5 subunits. The dominant factor that determines if a coiled-coil domain exists in solution as a homo or heterodimer is the charge placement of the amino acids present on the **e** and **g** positions of the heptad. The placement of surface charges is a critical component within the coiled-coil motif, which aids in directing complex formation and subsequent function (*13, 25-28*).

The coiled-coil structural motif has drawn considerable interest due to its diversity in structure, oligomerization state comprising 2 to 5 α-helical strands running parallel or anti-parallel, variations in length up to 1000 residues, use as a model system for protein design and as a tag system for affinity purification and biosensor applications (*22, 23, 29-31*). It is therefore relevant to design a coiled-coil system that is maximally stabilized for these applications. We have previously modified *de novo* coiled-coil sequences by varying parameters such as hydrophobicity, α-helical propensity, and chain length. The optimal combination of stability and specificity was observed with the sequence IAAL at positions **abcd** with a length of 21 residues (*32*). The heterodimeric peptides were designed in an E/K coil fashion where both subunits share the common IAAL repeat and the **efg** repeat for the E-coil contained **EKE** and the K-coil contained **KEK**. The resulting *de novo* designed system was therefore termed IAAL-E3/K3 coiled-coil. Complementary positioning of the charged residues between the E3/K3 coils ensured a maximal favorable electrostatic interaction. The resulting E3/K3 coiled-coil complex is a remarkably stable heterodimer, with a denaturation point of 4.3M GdnHCl, dissociation constant of 70*n*M and a ΔG^{H2O} = -9.6 kcal/mol (*32*).

Here we present the NMR solution structure of the IAAL-E3/K3 coiled-coil complex. The complex is shown to be an ideal coiled-coil domain, containing two parallel α-helices in a larger left-handed coiled-coil quaternary structure, consisting of a tightly packed hydrophobic core flanked on either side by charged residues. The positioning of Glu and Lys residues play an important role in heterodimer formation, eliminating any possible homodimeric formation for either the E3 or K3 subunits due to electrostatic repulsions. The IAAL-E3/K3 coiled-coil high-resolution NMR structure gives insights into future *de novo* protein designs, and is a prominent example of modern protein engineering using model systems to elucidate structural understandings and probe biological questions within nature.

EXPERIMENTAL PROCEDURES

F.1 Synthesis

Peptides IAAL-E3 and IAAL-K3 coils were synthesized by standard *t*-butyloxycarbonyl solid-phase techniques, purified, and characterized as previously described.(*32*) The synthetic peptide IAAL-E3 has the sequence Ac-EIAALEKEIAALEKEIAALEK-NH_2 and peptide IAAL-K3 has the sequence Ac-KIAALKEKIAALKEKIAALKE-NH_2.

F.2 NMR Spectroscopy

Lyophilized peptides IAAL-E3 and IAAL-K3 were dissolved in 500*u*L of 100mM KCl, 50mM KPO_4 in 90%H_2O / 10%D_2O buffer solution (pH 6.7) to a final concentration of 1.0mM each, with DSS added as an internal ^1H reference. Initially, one-dimensional ^1H-NMR was performed on the complex at 500MHz (sw=6500Hz, nt=32). Three subsequent 50*u*L additions of TFE-d_3 (Cambridge Isotope Laboratories) were added directly to the NMR tube, followed by two-dimensional NMR experiments for structural determination.

A natural abundance 2D {^1H, ^{15}N}-HSQC spectra was acquired at 800MHz (sw=8000Hz, sw1=2000Hz, nt=384, ni=64) followed by 2D ^1H-^1H homonuclear TOCSY and NOESY experiments for side-chain assignments and structural restraint information [TOCSY(NOESY) parameters: sw/sw1=8000Hz, nt/ni=32/512, mix=60ms(150ms)]. All 800MHz experiments were run using ProteinPack (Varian Associates) with data processing performed using NMRPipe and cross-peak assignment using NMRView (*33, 34*).

All NMR data used in this study were acquired at 20°C using Varian INOVA 500 MHz and INOVA 800 MHz spectrometers. Both spectrometers are equipped with triple resonance probes and Z-axis pulsed field gradients (XYZ gradients for the 800 MHz).

For the coiled-coil complex, the chemical shift assignments of the backbone and side chain atoms, as well NOE inter-proton distance restraints were determined using the two-dimensional NMR experiments described above.

F.3 Structure Determination

Structures of the IAAL-E3/K3 coiled-coil were generated using the software Cyana (http://www.cyana.gsc.riken.go.jp/Software/Cyana.html) and its integrated auto-assignment module Candid (*35-37*). The chemical shift identification tolerance and transpose errors were set to 0.015 ppm in both proton dimensions. Starting with an initial 99 distance restraints derived from manually assigned NOE cross-peaks in NMRView (*34*) a total number of 100 structures were generated per Candid round with 8000 steps in the Cyana annealing protocol. For each candid run, 38ϕ and 38φ TALOS (*38*) derived angle restraints were incorporated (with a minimum error set manually to 20°). The final number of both manual and Candid assigned NOE cross-peaks used in structural determination was 256. The chemical shift degeneracy of all Leu and Ile methyl protons impaired the unambiguous assignments of long range NOE's necessary to properly define the quaternary structure of the coiled-coil. To refine the position of the two helices relative to one another, a total number of 30 homology based long-range restraints involving Ile-Ile [3] and Leu-Leu [27] proton-proton distances (see *Table F-1*) derived from the NMR structure of c-Jun were used (*39*). The best 20 low 'target function value' structures of Candid's final round were kept as an ensemble representation of the family of generated structures.

F.4 Coordinates

The atomic coordinates for the IAAL-E3/K3 heterodimeric coiled-coil complex have been deposited in the RCSB Protein Data Bank (*40*) (1U0I) and the ^{1}H/^{15}N-NMR chemical shifts have been deposited in the BioMagResBank (accession number 6260).

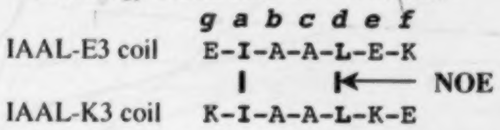
RESULTS AND DISCUSSION

Table F-1

Structural Statistics of the family of the 20 structures calculated

NOE restraints	
Manual & Candid assigned	256
Homology based	30
Dihedral restraints[a]	
ϕ	38
ψ	38
Backbone RMSD (Å)[b]	0.86
Restraint Violations[c]	
ϕ, ψ in core or allowed regions	100%

Homology Based Inter-coil restraints



E-coil	K-coil	Å[d]	E-coil	K-coil	Å[d]
L_{HB2}	L_{HB2}	2.44	L_{HD21}	L_{HB2}	3.79
L_{HD11}	L_{HB2}	4.57	L_{HB1}	L_{HB2}	3.63
L_{HG}	L_{HB2}	5.43	L_{HD21}	L_{HB1}	5.23
L_{HD21}	L_{HB1}	6.15	L_{HA}	L_{HB2}	4.77
L_{HN}	L_{HN}	7.90	I_{HG2}	I_{HG2}	3.20

[a] As calculated by TALOS software (*38*)

[b] Defined using residues 3-19 of both the E-coil and K-coil with respect to the average coordinates

[c] As determined using PROCHECK software (*41*)

[d] Proton-Proton distances as derived from the c-Jun coiled-coil homodimer (*39*)

There are numerous factors that must be considered when designing a heterodimeric coiled-coil protein system suitable for performing structural studies in solution. First there must be a binding constant of appropriate strength to allow for a single species to be present in solution. Second the protein system of interest must be adequately soluble in solution, and third the system must be of an appropriate molecular weight to allow for sufficiently narrow line widths for unambiguous NMR spectral assignments. Previously, it has been shown that the IAAL motif present in the heptad repeat, coupled with the advantageous placement of charged residues (Lys and Glu) for electrostatic interactions, is of sufficiently tight binding and solubility when comprised of only a mere three heptad repeats per subunit (*32*). This remarkably short, yet highly stable heterodimer is composed of only 21 residues per subunit, with the helical wheel diagram of proposed amino acid contacts displayed in Figure F-1. Peptides were additionally modified with an acetyl group on the N-termini and an amide group on the

C-termini to minimize any potential repulsive electrostatic interactions between the E3 and K3 subunits.

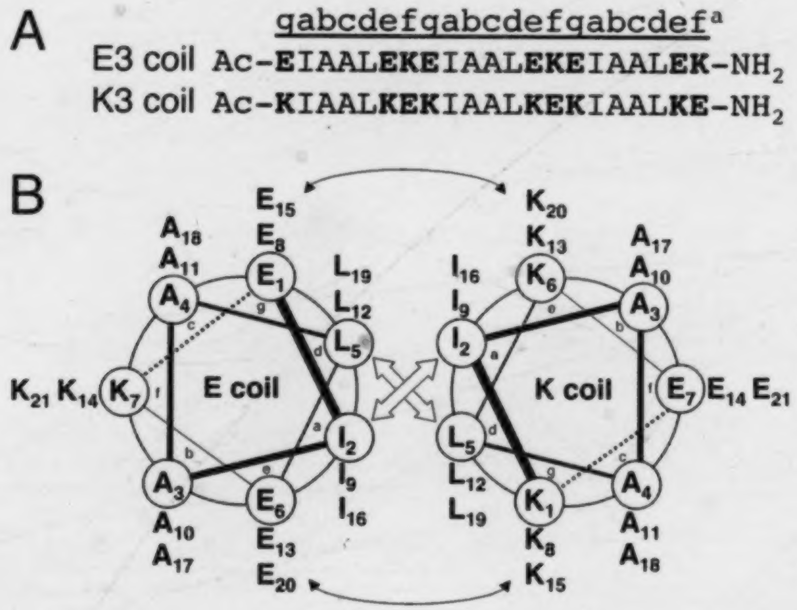

A
gabcdefgabcdefgabcdef a
E3 coil Ac-EIAALEKEIAALEKEIAALEK-NH2
K3 coil Ac-KIAALKEKIAALKEKIAALKE-NH2
B
E15
E8
E1
A18
A11
A4
L19
L12
L5
K21 K14 K7
E coil
I2
I9
I16
A3
A10
A17
E6
E13
E20
g
c
f
d
a
b
e
K20
K13
K6
I16
I9
I2
A17
A10
A3
K coil
E7 E14 E21
L5
L12
L19
K1
K8
K15
A4
A11
A18

Figure F-1. **Sequence and diagrammatic representation of the IAAL-E3/K3 protein complex**. (A) The four letter name (IAAL) denotes the peptide sequence in positions **a, b, c,** and **d** of the coiled-coil heptad repeat (^{a}*abcdefg*), and E and K (bolded) denotes the sequence variations between the E3 and K3 peptides in the **e, f** and **g** positions. Ac represents an N^{α}-acetyl group and NH_2 represents a C^{α}-amide group. (B) Helical wheel representation of the IAAL-E3/K3 coiled-coil heterodimer viewed in cross-section, as the peptide chain propagates into the page from the N- to the C-terminus. Predicted inter-helical hydrophobic interactions are denoted with wide arrows, and predicted inter-helical electrostatic interactions are denoted by the narrow arrows connecting the *i* to *i'*+5 positions.

The initial ^{1}H-NMR spectra acquired at 500MHz revealed a complex that was not homogeneous in solution. Figure F-2 shows the amide region of the ^{1}H-NMR spectrum, which displays a high degree of line broadening that can be attributed to unspecified quaternary interactions of the subunits forming a higher molecular weight multimer at the high concentrations used for NMR. Subsequent additions of 9.1%, 16.7% and 23.1%

(v/v) TFE-d_3 allowed for the breakup of all higher order molecular weight complexes to that of a single species in solution, as is indicated by the narrowing of lines within the ^{1}H-NMR spectrum. It has previously been shown that the CD ratios of $[\theta]_{222}/[\theta]_{208}$ for the IAAL-E3/K3 heterodimer have a ratio of 1.05 in 0% TFE and 0.83 in 50% TFE (*32*). Typically, values >1.0 are expected for interacting helices in a coiled-coil conformation and 0.85-0.95 for single α-helices (*42, 43*). The final concentration of 23.1% TFE was chosen to be sufficiently high to disrupt higher molecular weight complexes, yet low enough to prevent dissociation of the coiled-coil dimer to monomeric single-stranded α-helices. Although TFE is known to be a denaturant of quaternary structure when the concentrations of coiled-coils in solution are low as that required for CD analysis, the concentrations required for NMR are significantly higher and thus can maintain the monomer-dimer equilibrium in favor of dimer formation (*44*). However, the requirement of TFE to breakup aggregation and/or favor the dimer implies improvement to the robustness of the design may be possible (*45, 46*).

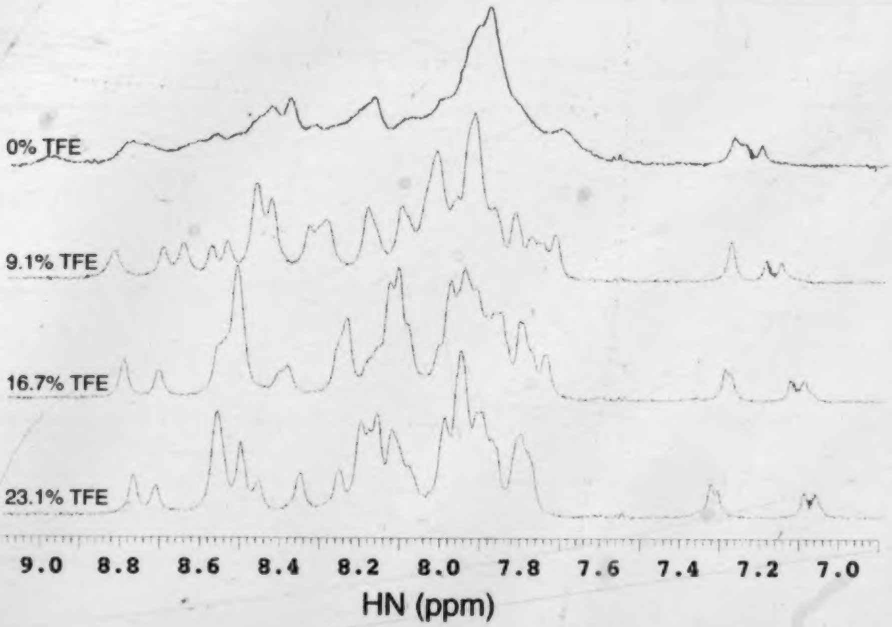


Figure F-2. **Amide proton resonance region of the ^{1}H-NMR spectrum.** ^{1}H-NMR spectra taken at various concentrations of TFE-d_3, showing progressive chemical shifts and line narrowing of backbone amide proton chemical shifts at 500MHz.

The stability of the E/K coil complex is well documented when used as an affinity purification system (*30*). It was shown to be stable to 0.5M NaCl and 80% acetonitrile washes. Because the coiled-coil heterodimerization domain is composed of and stabilized by both electrostatic and hydrophobic interactions at the dimer interface, high organic solvent (acetonitrile) would normally be predicted to break the non-covalent hydrophobic interactions between the two strands, however it actually stabilized the coiled-coil dimerization by enhancing electrostatic interactions between the strands (*47*). Similarly, high ionic strength salt that may disrupt the electrostatic interactions of the domain further enhanced the hydrophobic effect which occurs between the two α-helices. Thus, to disrupt the coiled-coil heterodimer both high concentration of organic solvent (50% acetonitrile) and a low pH of 2.0 were required. The low pH aided the protonation of the glutamic acid side-chains thereby disrupting the interhelical electrostatic interactions and the 50% acetonitrile disrupted the hydrophobic effect. Therefore it is reasonable that 23.1% TFE was unable to dissociate the heterodimer.

Upon finding a suitable buffer containing TFE for the production of a stable coiled-coil complex formation, a natural abundance 2D ^{15}N-HSQC was acquired at 800MHz to verify the presence of a single species in solution and demonstrate resolution, as shown in Figure F-3. The line-shape at 500MHz was insufficient to allow for well-resolved spectra, facilitating the need for higher field strengths (800MHz). Further acquisition of ^{1}H-TOCSY and ^{1}H-NOESY experiments allowed for the ^{1}H and ^{15}N assignments of backbone and side-chain resonances within the complex.

The repetitive nature of the heptad repeat comprising the sequences of the two subunits in the complex yielded a high degree of chemical shift redundancy. This is exemplified within the ^{15}N-HSQC spectra, regarding amide resonances such as $I9_E/I9_K$, $A11_E/A11_K$ and $E13_E/K13_K$. The majority of resonances within the ^{15}N-HSQC follow this overlapping pattern, revealing that resonances at the same sequence number within the two subunits all possess a relatively equal chemical environment. This proves true as well for the amide shifts for opposing electrostatic resonances such as $E20_E/K20_K$ and $E1_E/K1_K$. However, the chemical shift degeneracy of amide resonances is not surprising when considering the secondary structural characteristics, coupled with the potentially identical micro-environments within solution indicated by the helical wheel diagram in Figure F-1A.

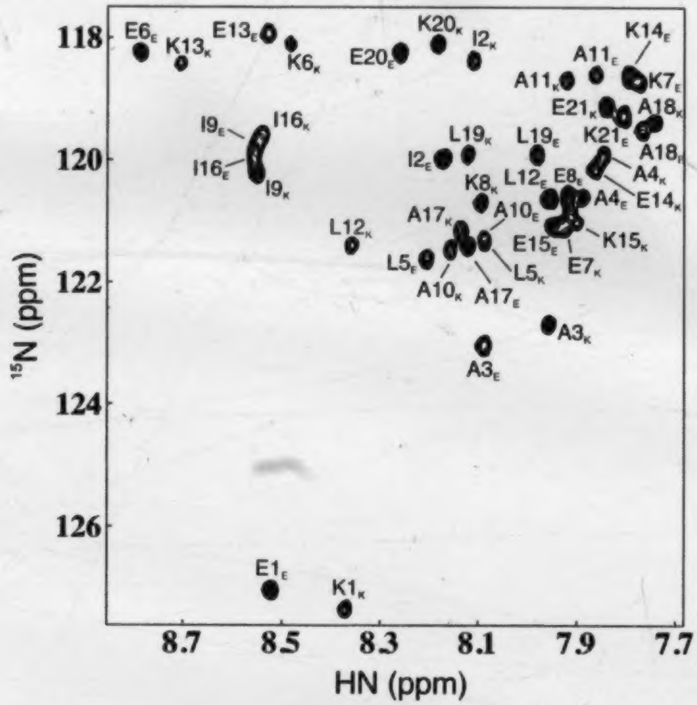


Figure F-3. **Amide proton resonance region of the ^{1}H-NMR spectrum.** 800MHz natural abundance ^{15}N-HSQC 2D-NMR spectra of the IAAL-E3/K3 coiled coil backbone amide groups at 20°C (23% TFE-d_3), with single amino acid letter representations for the amino acid cross-peaks and subscripts E and K denoting the cross-peak as originating from the E3 or the K3 coil in the complex.

The structure of the IAAL-E3/K3 coiled-coil is presented in Figure 4, shown in three different orientations of (A) an ensemble structure of the best 20 structures calculated, (B) a ribbon diagram of an averaged structure, and (C) a ribbon diagram looking down the helices. The ensemble family of structures reveals the super-imposition of the backbone heavy atoms, with the structural statistics of this ensemble reported in *Table F-1*. The two ribbon diagrams of the E3/K3 coiled-coil reveal two well-defined α-helices that overlap in a parallel fashion forming a left-handed super-coil, displaying the common quaternary fold indicative of this motif. The interhelical angle of the IAAL-E3/K3 coil-coil ensemble is 25 ± 3°, as evaluated from an in-house program (see *Appendix G* for complete structural statistics of 1U0I).

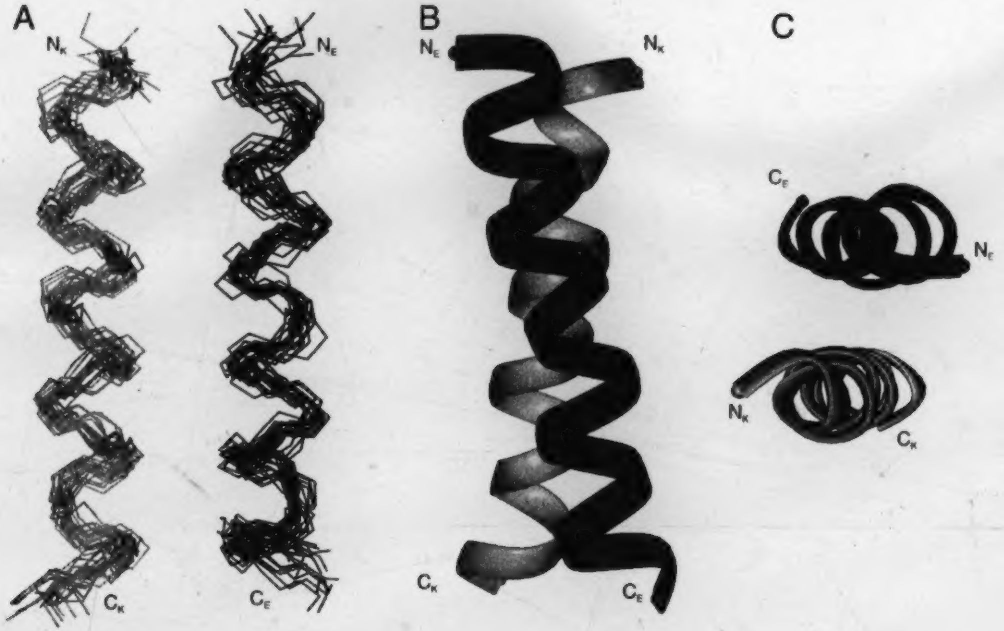


Figure F-4. **Structure of the heterodimeric IAAL-E3/K3 coiled-coil**. The E3/K3 coil is draw in dark/light with N-termini labeled as N_E/N_K and the C-termini labeled as C_E/C_K. (A) Backbone heavy atom super-imposition of the family of 20 generated structures of the IAAL-E3/K3 coiled-coil. (B) Ribbon diagram of the IAAL-E3/K3 coiled-coil complex rotated by $90°_y$. (C) Ribbon diagram of the IAAL-E3/K3 coiled-coil complex as the peptide chains propagate into the page from the N- to the C-terminus. The ensemble and ribbon diagrams were generated using DINO (*48*) and rendered using PovRay v3.5 for MacOSX (*49*).

The previously described amide chemical shift degeneracy was also observed in all side-chain groups containing CH_3 moieties. Due to the side-chain packing within the hydrophobic core of the complex, NOE cross-peaks for all Ile and Leu CH_3 groups present within the E3/K3 coils rendered all assignments completely ambiguous. This difficulty in resonance assignment rendered it impossible to determine if cross-peaks originated via *intra-* or *inter*-coil NOE contacts. It is however not uncommon to encounter chemical shift degeneracy when pursuing coiled-coil structures using NMR spectroscopy (*50*). Most coiled-coils that have been previously solved by NMR are biologically active complexes and although they possess the common heptad repeat, there is variation within the primary amino acid sequences of the subunits allowing for greater ease in assignment. The IAAL-E3/K3 coiled-coil was designed *de novo*, maximizing all positive stabilizing forces while minimizing any potential destabilizing influences (*32*). The perfectly conserved heptad repeats found in the E3/K3 coiled-coil negated any potentially helpful assignments due to amino acid sequence variation.

The ambiguity within the chemical shift resonances of the Leu and Ile residues required *inter*-coil NOE restraints to be added to the restraint file prior to the NMR simulated annealing protocol for structure determination, based upon homology modeling. The added NOE restraints can be viewed in *Table F-1* and are described in experimental procedures. These NOE restraints were derived from detailed analysis of the NMR solution structure of the c-Jun coiled-coil homodimer (*39*). The c-Jun homodimer was chosen as a 'model' coiled-coil system from which restraints could be derived due to its well-defined secondary and quaternary structure. As a result, the reported structure of the E3/K3 coiled-coil is similar to the c-jun homodimer in quaternary structure and implied hydrophobic contacts.

The packing of the amino acid side-chains within the E3/K3 coiled-coil is shown in Figure F-5. The hydrophobic residues Ile and Leu, as well as the electrostatic residues Lys and Glu, are displayed as ribbon diagrams to show the relative orientation and positioning of the individual amino acids. The stacking of the Leu and Ile residues are arranged as to create a tightly packed core within the protein, which is shown as a surface representation. The hydrophobic core is flanked on either side by an intricate network of oppositely charged Lys and Glu residues, positioned so as to form a crisscross pattern of charge across the hydrophobic interior.

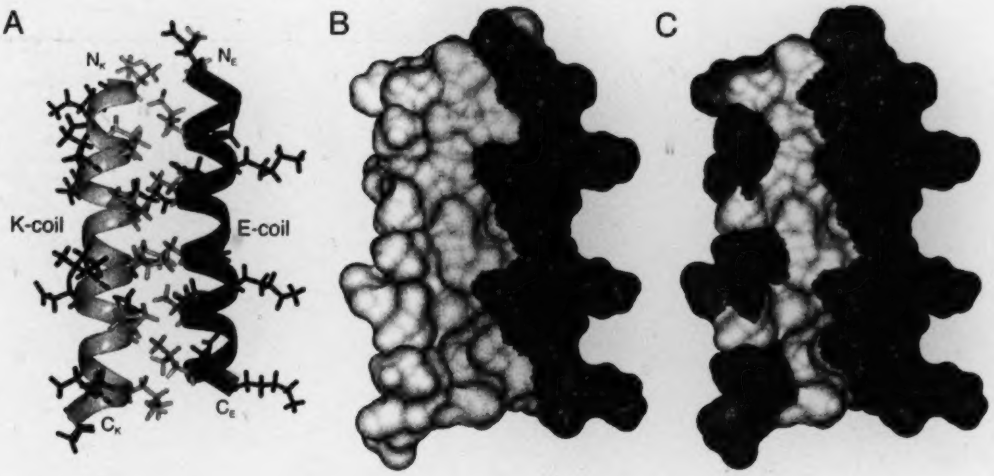


Figure F-5. **The structure of the heterodimeric IAAL-E3/K3 domain**. The E3/K3 coil is draw in black/white with N-termini labeled as N_E/N_K and the C-termini labeled as C_E/C_K. (A) Ribbon diagram displaying the amino acids which comprise the hydrophobic core are drawn in yellow (Ile and Leu) and the hydrophilic residues responsible for electrostatic interactions drawn in blue (Lys) and red (Glu). (B) Hydrophobic surface map representation of the IAAL-E3/K3 coiled-coil complex, revealing the hydrophobic core with Ile and Leu drawn in yellow. (C) Electrostatic surface map representation of the IAAL-E3/K3 coiled-coil complex, revealing the surface hydrophilic residues Glu (red) and Lys (blue). The ribbon diagram was generated using DINO (*48*) and rendered using PovRay v3.5 for MacOSX (*49*), and surface diagrams were generated using InsightII v98 (Accelrys Inc.).

The hydrophobic effect has been hypothesized to be the dominant driving force in protein folding, and hence has been the primary focus in understanding the directing effects of protein folding both *in vivo* and *in vitro* (*51-54*), yet electrostatic interactions remain a critical component of protein structure and stability. The proposed formation of salt-bridges between the Lys and Glu residues of the two subunits are traditionally difficult to measure by solution NMR primarily due to the large distances involved which can be greater than the NOE detection limit of approximately 5Å ($I_{NOE} \alpha 1/r^6$) and the short residence times of the exchangeable protons present on the carboxylate (COOH) and amine (NH_3^+) groups which are freely solvent accessible, rendering them non-observable during spectral acquisition as they resonate with the bulk water signal.

The contribution of charged/neutral residues along the periphery of coiled-coil systems to the stabilizing nature of electrostatic interactions has long been controversial (*55, 56*). It is known that the precise placement of charged residues is primarily important in directing heterodimeric versus homodimeric formation (*57, 58*). however the immediate stabilizing effects directing salt-bridge formation has been argued if they indeed form at all (*56*). It is known from crystallographic studies of coiled-coil systems that visualization of certain salt-bridges is realized, yet based upon sequence alignments not all predicted-salt bridges occur and are thus not observed (*11, 59*). This may lead one to question whether salt-bridge formation occurs at all in a solution environment when entropic considerations regarding desolvation penalties must be considered. However, it has previously been reported based upon double-mutant cycle analyses that the net electrostatic contribution per interhelical (*i* to *i*'+5) Glu-Lys salt-bridge is –0.4 kcal/mol (*27*), which is directly applicable to the IAAL-E3/K3 system. It is therefore our opinion that since the attractive forces between two charged particles occur over much larger distances than shorter-range hydrophobic interactions (>5Å), salt-bridges must be plausible on the periphery of the E3/K3 system and are aiding in directing the high stability properties of this complex.

The surface rendering of the IAAL-E3/K3 coiled-coil electrostatic residues in Figure F-5C shows the positioning of the Lys and Glu residues within the complex, oriented to allow for maximal attractive interactions between the two subunits. Based upon the positioning of the residues, it is visually clear how the charged residues may play a key role in directing a homo versus heterodimeric complex formation as the charge repulsion for homodimer formation of either subunit would be too great when compared to that for a heterodimer (*55*). However, we have previously shown that a *de novo* coiled-coil identical to the E3/K3 system with an additional heptad repeat (IAAL-E4/K4) displayed a small degree of homodimer formation, where the additional hydrophobic stability provided by the extra heptad repeat is predicted to partially compensate for the charge repulsions. However for the E3/K3 system, a reduced probability of homodimer formation is observed as the charge repulsions play a more dominant role over the stabilizing effects of the shorter number of IAAL repeats seen in the E4/K4 system (*32*). Therefore the perfectly designed complementarity of Lys and Glu between the two subunits is a key factor when initiating coiled-coil formation. If the entropic solvent effect of the hydrophobic residues within the subunits is sufficient for the two subunits to bind in a 'knobs into holes' fashion (*60*), then the interactions between the electrostatic residues are responsible for the specificity of heterodimer formation and the tight binding constant.

Table F-2

Distances of Salt-Bridging pairs within IAAL-E3/K3 coiled-coil

	Electrostatic Pair[a]	*Mean(Å)*[b]	*Lowest(Å)*	*Highest(Å)*
A	$E1_E$ — $K6_K$	14.1 ± 2.9	8.5	20.0
	$E8_E$ — $K13_K$	10.7 ± 3.3	6.4	18.7
	$E15_E$ — $K20_K$	10.5 ± 3.1	5.0	18.1
	$K1_K$ — $E6_E$	16.7 ± 1.8	13.3	20.1
	$K8_K$ — $E13_E$	10.9 ± 3.1	6.4	19.3
	$K15_K$ — $E20_E$	9.6 ± 3.0	6.1	17.8
B	$E6_E$ — $K8_K$	9.3 ± 2.5	3.6	15.4
	$E13_E$ — $K15_K$	9.5 ± 2.2	4.8	15.1
	$K6_K$ — $E8_E$	10.2 ± 2.3	4.0	13.4
	$K13_K$ — $E15_E$	11.4 ± 2.6	6.6	16.2

[a] Electrostatic pairs present on the coiled-coil surface where *i* to *i*'+5 interactions are denoted by **A**, and *i* to *i*'+2 by **B**

[b] Average value taken from the family of 20 structures calculated, with charged group centroids defined as the side-chain carboxylate carbon (CD) for Glu and amine nitrogen (NZ) for Lys residues.

The observed distances separating the putative salt-bridging pairs within the E3/K3 coiled-coil are listed in *Table F-2*. The average distances (excluding termini) for the *i* to *i*'+5 contacts range from 9.6Å to 10.9Å calculated over the family of 20 structures. Larger *i* to *i*'+5 distances of 14.1Å and 16.7Å were observed between the N-termini residues $E1_E$ and $K1_K$ and their respective pairing partners, which is attributed to the lower number of NOE restraints for the N-termini residues, yielding an overall larger RMSD for these regions. It is also noteworthy that the distances observed for the *i* to *i*'+2 interactions within the E3/K3 coiled-coil range from 9.3Å to 11.4Å, which are within the error ranges observed for the *i* to *i*'+5 interactions. We propose that the IAAL-E3/K3 coiled-coil is additionally stabilized by electrostatic interactions in an *i* to *i*'+2 fashion, in addition to the commonly reported *i* to *i*'+5 fashion. The distances derived for the E3/K3 coiled-coil are in close agreement with the NMR structure distances observed for the c-jun homodimer for both the *i* to *i*'+2 and the *i* to *i*'+5 electrostatic pairings (*39*).

The NMR solution structure of a highly stable engineered coiled-coil complex is presented. The IAAL-E3/K3 heterodimer provides a structural framework for current *de novo* design, which may subsequently lead to increased commercial availability of coiled-coil use within affinity purification and other biotechnical applications.

REFERENCES

1. Lavigne, P., Kondejewski, L. H., Houston, M. E., Jr., Sonnichsen, F. D., Lix, B., Skyes, B. D., Hodges, R. S., and Kay, C. M. (1995) *J Mol Biol 254*, 505-20.
2. Jean-Francois, N., Frederic, G., Raymund, W., Benoit, C., and Lavigne, P. (2003) *J Mol Biol 326*, 1577-95.
3. Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U., and Moarefi, I. (2000) *Cell 103*, 621-32.
4. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) *Nature 424*, 35-41.
5. Lupas, A. (1996) *Trends Biochem Sci 21*, 375-82.
6. Kohn, W. D., Mant, C. T., and Hodges, R. S. (1997) *J Biol Chem 272*, 2583-6.
7. Burkhard, P., Stetefeld, J., and Strelkov, S. V. (2001) *Trends Cell Biol 11*, 82-8.
8. Singh, M., Berger, B., and Kim, P. S. (1999) *J Mol Biol 290*, 1031-41.
9. Malashkevich, V. N., Singh, M., and Kim, P. S. (2001) *Proc Natl Acad Sci U S A 98*, 8502-6.
10. Cupp-Vickery, J. R., and Vickery, L. E. (2000) *J Mol Biol 304*, 835-45.
11. Glover, J. N., and Harrison, S. C. (1995) *Nature 373*, 257-61.
12. Liu, J., and Rost, B. (2001) *Protein Sci 10*, 1970-9.
13. Kohn, W. D., and Hodges, R. S. (1998) *Trends Biotechnol 16*, 379-389.
14. Muller, K. M., Arndt, K. M., and Alber, T. (2000) *Methods Enzymol 328*, 261-82.
15. Walshaw, J., and Woolfson, D. N. (2003) *J Struct Biol 144*, 349-61.
16. Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R. S. (1999) *Protein Sci 8*, 2312-29.
17. Zhu, B. Y., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *Protein Sci 2*, 383-94.
18. Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T., and Hodges, R. S. (2000) *J Mol Biol 300*, 377-402.
19. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) *Science 262*, 1401-7.
20. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature 371*, 80-3.
21. Woolfson, D. N., and Alber, T. (1995) *Protein Sci 4*, 1596-607.
22. Chao, H., Bautista, D. L., Litowski, J., Irvin, R. T., and Hodges, R. S. (1998) *J Chromatogr B Biomed Sci Appl 715*, 307-29.
23. Chao, H., Houston, M. E., Jr., Grothe, S., Kay, C. M., O'Connor-McCourt, M., Irvin, R. T., and Hodges, R. S. (1996) *Biochemistry 35*, 12175-85.
24. O'Neil, K. T., and DeGrado, W. F. (1990) *Science 250*, 646-51.
25. Krylov, D., Mikhailenko, I., and Vinson, C. (1994) *Embo J 13*, 2849-61.
26. Vinson, C. R., Hai, T., and Boyd, S. M. (1993) *Genes Dev 7*, 1047-58.
27. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1994) *Protein Eng 7*, 1365-72.
28. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1994) *J Mol Biol 237*, 500-12.
29. Catimel, B., Faux, M. C., Nerrie, M., Rothacker, J., Otvos, L. J., Wade, J. D., Nice, E. C., and Burgess, A. W. (2001) *J Pept Res 58*, 493-503.
30. Tripet, B., Yu, L., Bautista, D. L., Wong, W. Y., Irvin, R. T., and Hodges, R. S. (1996) *Protein Eng 9*, 1029-42.
31. Wolber, V., Maeda, K., Schumann, R., Brandmeier, B., Wiesmuller, L., and Wittinghofer, A. (1992) *Biotechnology (N Y) 10*, 900-4.

32. Litowski, J. R., and Hodges, R. S. (2002) *J Biol Chem 277*, 37272-9.
33. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.
34. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603-614.
35. Guntert, P. (2003) *Prog. NMR Spectrosc. 43*, 105-125.
36. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J Mol Biol 273*, 283-98.
37. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) *J Biomol NMR 24*, 171-89.
38. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J Biomol NMR 13*, 289-302.
39. Junius, F. K., O'Donoghue, S. I., Nilges, M., Weiss, A. S., and King, G. F. (1996) *J Biol Chem 271*, 13663-7.
40. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourné, P. E. (2000) *Nucleic Acids Res 28*, 235-42.
41. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thorton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-290.
42. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1992) *J Biol Chem 267*, 2664-70.
43. Lau, S. Y., Taneja, A. K., and Hodges, R. S. (1984) *J Biol Chem 259*, 13253-61.
44. Slupsky, C. M., Reinach, F. C., Smillie, L. B., and Sykes, B. D. (1995) *Protein Sci 4*, 1279-90.
45. Burkhard, P., Ivaninskii, S., and Lustig, A. (2002) *J Mol Biol 318*, 901-10.
46. Meier, M., Lustig, A., Aebi, U., and Burkhard, P. (2002) *J Struct Biol 137*, 65-72.
47. Lau, S. Y., Taneja, A. K., and Hodges, R. S. (1985) *Journal of Chromatography A 317*, 129-140.
48. DINO. (2002) *Visualizing Structural Biology*, http://www.dino3d.org.
49. PovRay. (2004) *Persistence of Vision Ray-Tracer*, http://www.povray.org.
50. O'Donoghue, S. I., and Nigles, M. (1999) in *Biological Magnetic Resonance, Volume 17: Structure Computation and Dynamics in Protein NMR* (Berliner, K., Ed.) pp 131-161, Kluwer Academic/Plenum Publishers, New York.
51. Kovacs, H., Mark, A. E., and van Gunsteren, W. F. (1997) *Proteins 27*, 395-404.
52. Tsai, C. J., Lin, S. L., Wolfson, H. J., and Nussinov, R. (1997) *Protein Sci 6*, 53-64.
53. Walls, P. H., and Sternberg, M. J. (1992) *J Mol Biol 228*, 277-97.
54. Dill, K. A. (1990) *Biochemistry 29*, 7133-55.
55. Lavigne, P., Sonnichsen, F. D., Kay, C. M., and Hodges, R. S. (1996) *Science 271*, 1136-8.
56. Lumb, K. J., and Kim, P. S. (1995) *Science 268*, 436-9.
57. Monera, O. D., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *J Biol Chem 268*, 19218-27.
58. Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) *Biochemistry 33*, 3862-71.
59. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science 254*, 539-44.
60. Crick, F. H. C. (1953) *Acta Crystallogr 6*, 689-698.

APPENDIX G

NMR solution structure of a highly stable de novo heterodimeric coiled-coil – chemical shift and procheck structural analysis

This appendix contains additional information regarding structural characterization of the presented NMR solution structure of the EK heterodimeric coiled-coil described in *Appendix F* of this thesis. This work was published in the December 2004 issue of the journal *Biopolymers* [*Biopolymers*. 2004 Dec 5;**75(5)**:367-7.]. All chemical shift data are reported as standard output (*.str) from the BioMagResBank (BMRB accession code 6260) and do not reflect the precision of the chemical shifts.

Appendix G-1. Assigned ^{1}H and ^{15}N chemical shifts of EK coiled-coil

A. E-coil assigned chemical shifts

1	1	GLU	N	N	127.040	0.05	1
2	1	GLU	H	H	8.525	0.02	1
3	1	GLU	HA	H	4.221	0.02	1
4	1	GLU	HB3	H	2.092	0.02	2
5	1	GLU	HB2	H	2.021	0.02	2
6	1	GLU	HG3	H	2.355	0.02	2
7	2	ILE	N	N	119.971	0.05	1
8	2	ILE	H	H	8.170	0.02	1
9	2	ILE	HA	H	3.953	0.02	1
10	2	ILE	HB	H	1.954	0.02	1
11	2	ILE	HG12	H	1.301	0.02	1
12	2	ILE	HD1	H	0.949	0.02	1
13	2	ILE	HG2	H	0.910	0.02	1
14	3	ALA	N	N	123.043	0.05	1
15	3	ALA	H	H	8.085	0.02	1
16	3	ALA	HA	H	4.260	0.02	1
17	3	ALA	HB	H	1.557	0.02	1
18	4	ALA	N	N	120.598	0.05	1
19	4	ALA	H	H	7.887	0.02	1
20	4	ALA	HA	H	4.187	0.02	1
21	4	ALA	HB	H	1.611	0.02	1
22	5	LEU	N	N	121.626	0.05	1
23	5	LEU	H	H	8.207	0.02	1
24	5	LEU	HA	H	4.184	0.02	1
25	5	LEU	HB3	H	1.892	0.02	2
26	5	LEU	HB2	H	1.813	0.02	2
27	5	LEU	HD1	H	0.924	0.02	2
28	5	LEU	HD2	H	0.961	0.02	2
29	6	GLU	N	N	118.245	0.05	1
30	6	GLU	H	H	8.786	0.02	1
31	6	GLU	HA	H	3.887	0.02	1
32	6	GLU	HB3	H	2.071	0.02	2
33	6	GLU	HB2	H	2.180	0.02	2
34	6	GLU	HG3	H	2.246	0.02	2
35	6	GLU	HG2	H	2.620	0.02	2
36	7	LYS	N	N	118.720	0.05	1
37	7	LYS	H	H	7.770	0.02	1
38	7	LYS	HA	H	4.113	0.02	1
39	7	LYS	HB3	H	2.037	0.02	2
40	7	LYS	HG2	H	1.514	0.02	2
41	7	LYS	HD2	H	1.693	0.02	2
42	7	LYS	HE2	H	3.019	0.02	2
43	8	GLU	N	N	120.581	0.05	1
44	8	GLU	H	H	7.911	0.02	1
45	8	GLU	HA	H	4.184	0.02	1
46	8	GLU	HB3	H	2.216	0.02	2
47	8	GLU	HB2	H	2.276	0.02	2
48	8	GLU	HG3	H	2.411	0.02	2
49	9	ILE	N	N	119.874	0.05	1
50	9	ILE	H	H	8.554	0.02	1
51	9	ILE	HA	H	3.689	0.02	1
52	9	ILE	HB	H	1.945	0.02	1
53	9	ILE	HG13	H	1.106	0.02	1
54	9	ILE	HG12	H	0.949	0.02	1
55	9	ILE	HD1	H	0.922	0.02	1
56	9	ILE	HG2	H	0.842	0.02	1
57	10	ALA	N	N	121.327	0.05	1
58	10	ALA	H	H	8.086	0.02	1
59	10	ALA	HA	H	4.260	0.02	1
60	10	ALA	HB	H	1.480	0.02	1
61	11	ALA	N	N	118.591	0.05	1
62	11	ALA	H	H	7.858	0.02	1
63	11	ALA	HA	H	4.198	0.02	1
64	11	ALA	HB	H	1.548	0.02	1
65	12	LEU	N	N	120.643	0.05	1
66	12	LEU	H	H	7.954	0.02	1
67	12	LEU	HA	H	4.205	0.02	1
68	12	LEU	HB3	H	1.806	0.02	2
69	13	GLU	N	N	117.927	0.05	1
70	13	GLU	H	H	8.529	0.02	1

71	13	GLU	HA	H	3.895	0.02	1
72	13	GLU	HB3	H	2.208	0.02	2
73	13	GLU	HB2	H	2.075	0.02	2
74	13	GLU	HG3	H	2.545	0.02	2
75	13	GLU	HG2	H	2.620	0.02	2
76	14	LYS	N	N	118.630	0.05	1
77	14	LYS	H	H	7.791	0.02	1
78	14	LYS	HA	H	4.121	0.02	1
79	14	LYS	HE2	H	3.020	0.02	2
80	15	GLU	N	N	121.077	0.05	1
81	15	GLU	H	H	7.945	0.02	1
82	15	GLU	HA	H	4.198	0.02	1
83	15	GLU	HB3	H	2.203	0.02	2
84	15	GLU	HB2	H	2.269	0.02	2
85	15	GLU	HG3	H	2.400	0.02	2
86	16	ILE	N	N	119.874	0.05	1
87	16	ILE	H	H	8.557	0.02	1
88	16	ILE	HA	H	3.687	0.02	1
89	16	ILE	HB	H	1.944	0.02	1
90	17	ALA	N	N	121.424	0.05	1
91	17	ALA	H	H	8.120	0.02	1
92	17	ALA	HA	H	4.181	0.02	1
93	17	ALA	HB	H	1.533	0.02	1
94	18	ALA	N	N	119.388	0.05	1
95	18	ALA	H	H	7.739	0.02	1
96	18	ALA	HA	H	4.186	0.02	1
97	18	ALA	HB	H	1.569	0.02	1
98	19	LEU	N	N	119.903	0.05	1
99	19	LEU	H	H	7.980	0.02	1
100	19	LEU	HA	H	4.182	0.02	1
101	19	LEU	HB3	H	2.005	0.02	2
102	19	LEU	HG	H	1.641	0.02	1
103	19	LEU	HD1	H	0.953	0.02	2
104	20	GLU	N	N	118.263	0.05	1
105	20	GLU	H	H	8.260	0.02	1
106	20	GLU	HA	H	4.124	0.02	1
107	20	GLU	HB3	H	2.121	0.02	2
108	20	GLU	HB2	H	2.260	0.02	2
109	20	GLU	HG3	H	2.499	0.02	2
110	20	GLU	HG2	H	2.520	0.02	2
111	21	LYS	N	N	119.269	0.05	1
112	21	LYS	H	H	7.806	0.02	1
113	21	LYS	HA	H	4.235	0.02	1
114	21	LYS	HB3	H	1.977	0.02	2
115	21	LYS	HB2	H	1.929	0.02	2

B. K-coil assigned chemical shifts

1	1	LYS	N	N	127.345	0.05	1
2	1	LYS	H	H	8.370	0.02	1
3	1	LYS	HA	H	4.137	0.02	1
4	1	LYS	HB3	H	1.863	0.02	2
5	1	LYS	HB2	H	2.159	0.02	2
6	1	LYS	HG3	H	1.485	0.02	2
7	1	LYS	HG2	H	1.593	0.02	2
8	1	LYS	HD2	H	1.758	0.02	2
9	1	LYS	HE2	H	3.014	0.02	2
10	2	ILE	N	N	118.369	0.05	1
11	2	ILE	H	H	8.108	0.02	1
12	2	ILE	HA	H	3.894	0.02	1
13	2	ILE	HB	H	1.991	0.02	1
14	2	ILE	HG13	H	1.307	0.02	1
15	2	ILE	HG12	H	1.483	0.02	1
16	2	ILE	HG2	H	0.955	0.02	1
17	3	ALA	N	N	122.701	0.05	1
18	3	ALA	H	H	7.953	0.02	1
19	3	ALA	HA	H	4.057	0.02	1
20	3	ALA	HB	H	1.489	0.02	1
21	4	ALA	N	N	119.926	0.05	1
22	4	ALA	H	H	7.846	0.02	1
23	4	ALA	HA	H	4.229	0.02	1
24	4	ALA	HB	H	1.535	0.02	1
25	5	LEU	N	N	121.318	0.05	1
26	5	LEU	H	H	8.092	0.02	1
27	5	LEU	HA	H	4.148	0.02	1
28	5	LEU	HB3	H	1.798	0.02	2
29	5	LEU	HB2	H	1.848	0.02	2
30	5	LEU	HD2	H	0.923	0.02	2
31	6	LYS	N	N	118.096	0.05	1
32	6	LYS	H	H	8.483	0.02	1
33	6	LYS	HA	H	3.900	0.02	1
34	6	LYS	HB3	H	1.926	0.02	2
35	6	LYS	HB2	H	1.970	0.02	2
36	6	LYS	HG2	H	1.395	0.02	2
37	7	GLU	N	N	120.930	0.05	1
38	7	GLU	H	H	7.918	0.02	1
39	7	GLU	HA	H	4.122	0.02	1
40	8	LYS	N	N	120.696	0.05	1
41	8	LYS	H	H	8.086	0.02	1
42	8	LYS	HA	H	4.155	0.02	1
43	8	LYS	HB3	H	1.798	0.02	2
44	8	LYS	HB2	H	1.848	0.02	2
45	9	ILE	N	N	120.203	0.05	1
46	9	ILE	H	H	8.551	0.02	1
47	9	ILE	HA	H	3.689	0.02	1
48	9	ILE	HB	H	1.945	0.02	1
49	9	ILE	HD1	H	0.845	0.02	1
50	9	ILE	HG2	H	0.931	0.02	1
51	10	ALA	N	N	121.490	0.05	1
52	10	ALA	H	H	8.157	0.02	1
53	10	ALA	HA	H	4.064	0.02	1
54	10	ALA	HB	H	1.574	0.02	1
55	11	ALA	N	N	118.673	0.05	1
56	11	ALA	H	H	7.922	0.02	1
57	11	ALA	HA	H	4.232	0.02	1
58	11	ALA	HB	H	1.601	0.02	1
59	12	LEU	N	N	121.397	0.05	1
60	12	LEU	H	H	8.357	0.02	1
61	12	LEU	HA	H	4.144	0.02	1
62	12	LEU	HB3	H	1.918	0.02	2
63	12	LEU	HG	H	1.752	0.02	1
64	12	LEU	HD1	H	0.917	0.02	2
65	12	LEU	HD2	H	0.947	0.02	2
66	13	LYS	N	N	118.427	0.05	1
67	13	LYS	H	H	8.705	0.02	1
68	13	LYS	HA	H	3.896	0.02	1
69	13	LYS	HB3	H	1.989	0.02	2
70	13	LYS	HB2	H	1.922	0.02	2
71	13	LYS	HG2	H	1.382	0.02	2
72	13	LYS	HE2	H	2.887	0.02	2
73	14	GLU	N	N	120.122	0.05	1
74	14	GLU	H	H	7.859	0.02	1
75	14	GLU	HA	H	4.108	0.02	1
76	14	GLU	HB3	H	2.210	0.02	2
77	14	GLU	HB2	H	2.104	0.02	2
78	14	GLU	HG3	H	2.337	0.02	2
79	14	GLU	HG2	H	2.498	0.02	2
80	15	LYS	N	N	121.003	0.05	1
81	15	LYS	H	H	7.896	0.02	1
82	15	LYS	HA	H	4.156	0.02	1

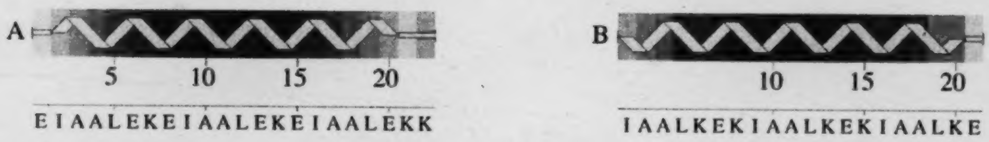
83	15	LYS	HB3	H	2.040	0.02	2
84	15	LYS	HB2	H	2.110	0.02	2
85	16	ILE	N	N	119.615	0.05	1
86	16	ILE	H	H	8.538	0.02	1
87	16	ILE	HA	H	3.661	0.02	1
88	16	ILE	HB	H	1.938	0.02	1
89	16	ILE	HG2	H	0.844	0.02	1
90	17	ALA	N	N	121.147	0.05	1
91	17	ALA	H	H	8.131	0.02	1
92	17	ALA	HA	H	4.050	0.02	1
93	17	ALA	HB	H	1.538	0.02	1
94	18	ALA	N	N	119.503	0.05	1
95	18	ALA	H	H	7.764	0.02	1
96	18	ALA	HA	H	4.215	0.02	1
97	18	ALA	HB	H	1.580	0.02	1
98	19	LEU	N	N	119.909	0.05	1
99	19	LEU	H	H	8.122	0.02	1
100	19	LEU	HA	H	4.166	0.02	1
101	19	LEU	HB3	H	2.017	0.02	2
102	19	LEU	HD1	H	0.916	0.02	2
103	20	LYS	N	N	118.088	0.05	1
104	20	LYS	H	H	8.181	0.02	1
105	20	LYS	HA	H	4.099	0.02	1
106	20	LYS	HB3	H	1.936	0.02	2
107	21	GLU	N	N	119.118	0.05	1
108	21	GLU	H	H	7.845	0.02	1
109	21	GLU	HA	H	4.211	0.02	1
110	21	GLU	HB3	H	2.103	0.02	2
111	21	GLU	HB2	H	2.176	0.02	2
112	21	GLU	HG3	H	2.362	0.02	2
113	21	GLU	HG2	H	2.448	0.02	2

Appendix G-2. Procheck analysis of EK coiled-coil

The following pages contain analysis of the NMR ensemble structure of the heterodimeric coiled-coil, as per the pdb coordinates **1U0I** from the protein data bank. Within the pdb file, the E-coil is defined as fragment A and the K-coil is defined as fragment B. This analysis includes all 20 deposited lowest energy structures that were deposited in the pdb. Complete Ramachandran plot analysis for EK coiled-coil and individual residues are reported, along with an example of predicted secondary structural characteristics as per procheck-NMR.

Model secondary structure

1U0I

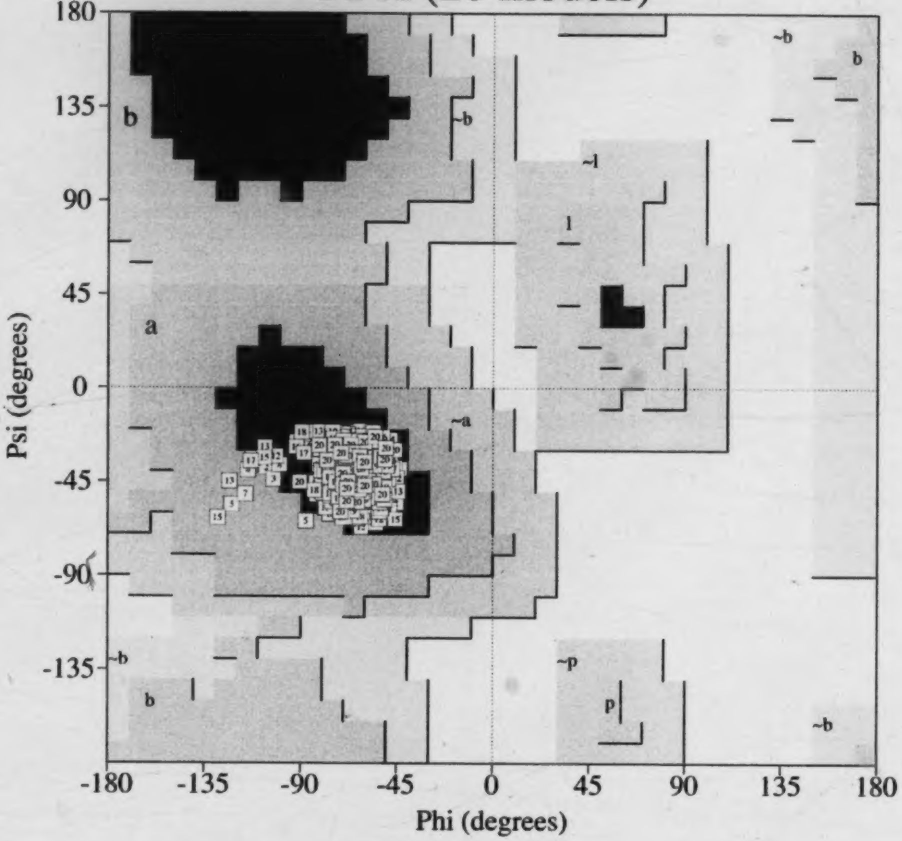

A
5
10
15
20
E I A A L E K E I A A L E K E I A A L E K K
B
10
15
20
I A A L K E K I A A L K E K I A A L K E

E-coil residues

K-coil residues

Ramachandran Plot

1U0I (20 models)



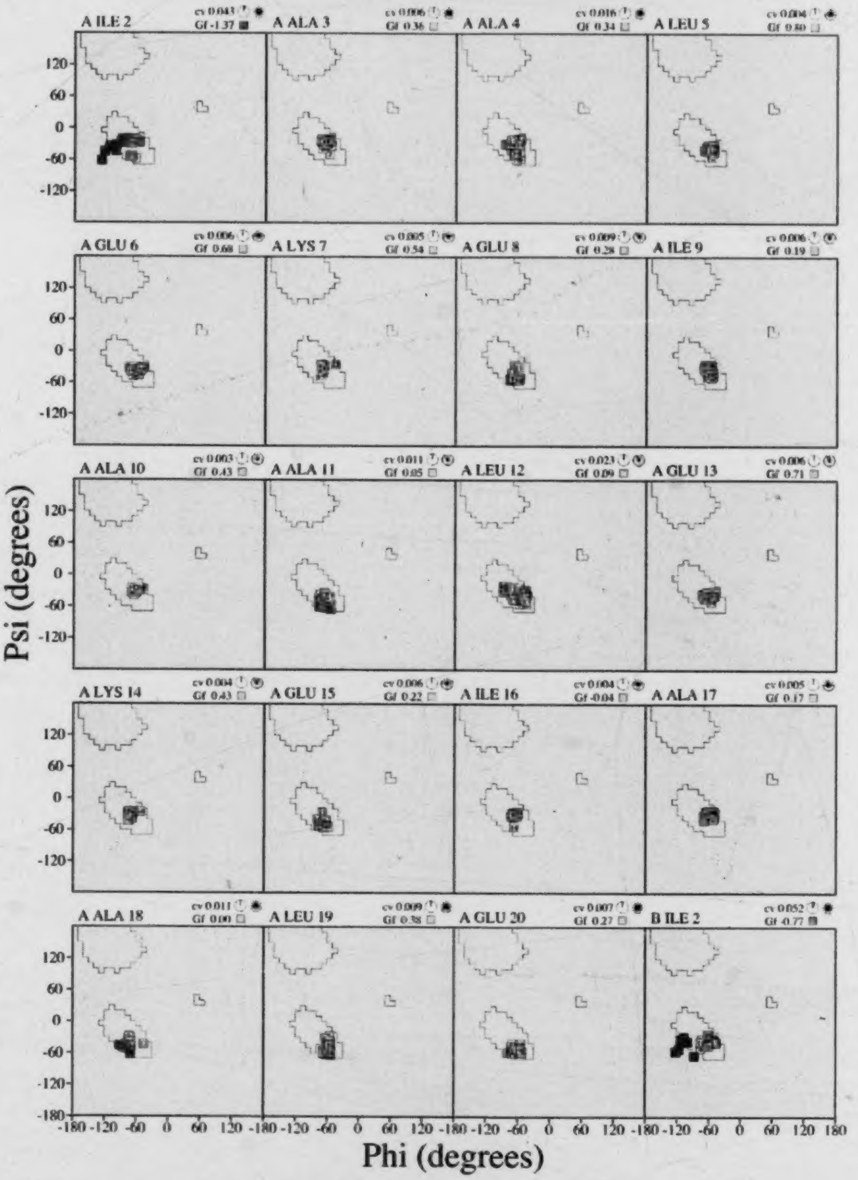
Plot statistics

Residues in most favoured regions [A,B,L]	749	98.6%
Residues in additional allowed regions [a,b,l,p]	11	1.4%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
	----	------
Number of non-glycine and non-proline residues	760	100.0%
Number of end-residues (excl. Gly and Pro)	80	
Number of glycine residues (shown as triangles)	0	
Number of proline residues	0	

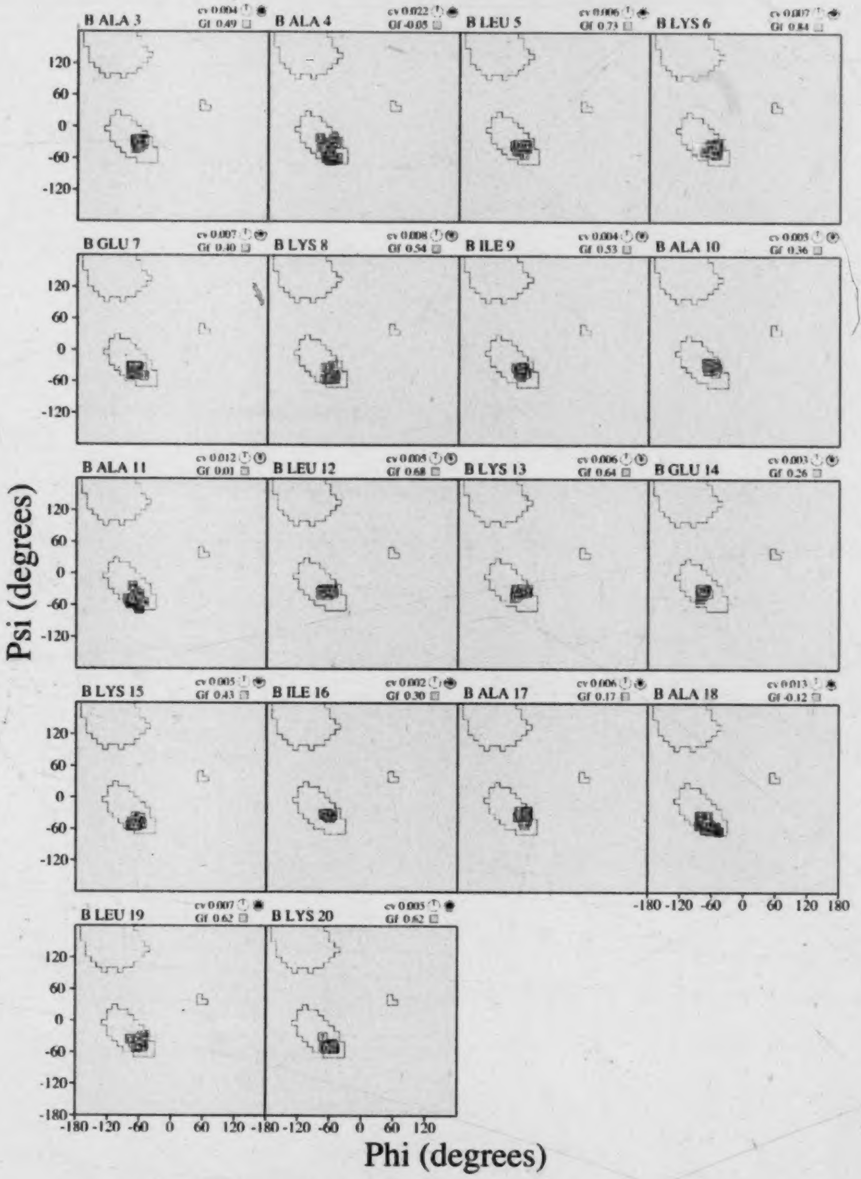
Total number of residues	840	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms
and R-factor no greater than 20%, a good quality model would be expected
to have over 90% in the most favoured regions.
Model numbers shown inside each data point.

Ensemble Ramachandran plots
1U0I (20 models)



cv = Circular Variance (low values signify high clustering of the data points). ● Accessible ⓘ Buried
Gf = Average G-factor for the residue (the higher the value the more favourable the conformations) based on analysis of high-res. Xstal structures
Data points coloured according to G-factor: Unfavourable



Ensemble Ramachandran plots
1U0I (20 models)

cv = Circular Variance (low values signify high clustering of the data points). ● Accessible ⦿ Buried
Gf = Average G factor for the residue (the higher the value the more favourable the conformations) based on analysis of high-res. Xstal structures
Data points coloured according to G-factor: Unfavourable